

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 15, 2004, 15:46:56 ; Search time 7228 Seconds
(without alignments)
4113.631 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618

Sequence: 1 MAPITAYAAQTRGLGLCIIT.....PAIPDREVLVREFDEMEEC 686

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US09930591/runat_15092004_164646_7096/app.query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09930591 @CGN_1_1_4545 @runat_15092004_164646_7096 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sv.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3618	100.0	2061	6	AX441176	Sequence AX441176
2	3618	100.0	2061	6	AX467113	Sequence AX467113
3	3574	98.8	5360	6	I06434	Sequence I06434
4	3574	98.8	5360	6	I09328	Sequence I09328
5	3574	98.8	6785	6	I06440	Sequence I06440
6	3574	98.8	6785	6	I09329	Sequence I09329
7	3574	98.8	7310	6	AX118696	Sequence AX118696
8	3574	98.8	7310	6	I09331	Sequence I09331
9	3574	98.8	7310	14	HPCPOLXP	Sequence HPCPOLXP
10	3574	98.8	9185	6	I08294	Sequence I08294
11	3574	98.8	9185	6	BD091382	Sequence BD091382
12	3574	98.8	9379	6	AX166930	Sequence AX166930
13	3574	98.8	9379	6	AX301300	Sequence AX301300
14	3574	98.8	9401	6	AX176483	Sequence AX176483
15	3574	98.8	9401	6	E66593	Sequence E66593
16	3574	98.8	9401	6	I71894	Sequence I71894
17	3574	98.8	9401	6	I81885	Sequence I81885
18	3574	98.8	9401	6	BD080334	Sequence BD080334
19	3574	98.8	9401	14	HPCPLYPRE	Sequence HPCPLYPRE
20	3574	98.8	9609	12	AF387805	Sequence AF387805
21	3574	98.8	9646	12	AF387806	Sequence AF387806
22	3574	98.8	9693	12	AF387807	Sequence AF387807
23	3571	98.7	6785	6	AX118692	Sequence AX118692
24	3570	98.7	8316	6	AX118703	Sequence AX118703
25	3570	98.7	9185	6	AX118722	Sequence AX118722
26	3570	98.7	9185	6	AX118723	Sequence AX118723
27	3570	98.7	9618	14	AF271632	Sequence AF271632
28	3568	98.6	6299	6	AX164584	Sequence AX164584
29	3565	98.5	2058	6	AX404933	Sequence AX404933
30	3565	98.5	2058	6	AX408362	Sequence AX408362
31	3565	98.5	2058	6	AX395309	Sequence AX395309
32	3565	98.5	2058	6	AX454818	Sequence AX454818
33	3565	98.5	8987	6	AX118728	Sequence AX118728
34	3565	98.5	9609	12	AF387808	Sequence AF387808
35	3555	98.3	5360	6	AX118686	Sequence AX118686
36	3551	98.1	9646	6	AX110828	Sequence AX110828
37	3551	98.1	9646	6	BD069982	Sequence BD069982
38	3551	98.1	12980	14	AF009606	Sequence AF009606
39	3551	98.1	12980	6	AX110831	Sequence AX110831
40	3551	98.1	12980	6	BD069985	Sequence BD069985
41	3550	98.1	9395	14	AF511950	Sequence AF511950
42	3550	98.1	9502	6	E08263	Sequence E08263
43	3550	98.1	9502	6	E08264	Sequence E08264
44	3550	98.1	9502	14	HPCHCJ1	Sequence HPCHCJ1
45	3547	98.0	9610	14	HEC278830	Sequence HEC278830

ALIGNMENTS

RESULT 1

AX441176	AX441176	2061 bp	DNA	linear	PAT 28-JUN-2002	
LOCUS	Sequence 16 from Patent WO0213855.					
DEFINITION	AX441176					
ACCESSION	AX441176					
VERSION	AX441176.1	GI:21665758				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	artificial sequences.					
REFERENCE	1					
AUTHORS	Sallberg, M. and Huitgren, C.					
TITLE	Vaccines containing ribavirin and methods of use thereof					
JOURNAL	Patent: WO 0213855-A 16 21-FEB-2002;					
	TRIPEP AB (SE)					
FEATURES	Location/Qualifiers					
source	1..2061					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
ORIGIN	/note="Hepatitis C virus NS3/4A coding region"					
Alignment Scores:						
Pred. No.:	1.41e-203	Length:	2061			
Score:	3618.00	Matches:	686			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			
US-09-930-591-2 (1-686) x AX441176 (1-2061)						
Qy	1	MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr	20			
Db	1	ATGGCGCTATCAGGGCTATGCCACACAGAGGGGCGCTTTGGGATGCATAATCACC	60			
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40			
Db	61	AGCTTGACCGCGCGGCAAAACCAGGTGAGGTTCAGATCGTGTCATCGCT	120			
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla	60			
Db	121	GCCACAGACTTCTGGCAACTGCATTAAACGGGGTGTGTGGACTGTACCATGGAGCC	180			
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnVetTyrThrAsnValAsp	80			
Db	181	GGAAACAGGACCAATGGCGTCACTTAAGGTCCTGTATCCAGATGTACACCAATGTGGAC	240			
Qy	81	GlnAspLeuValGlyTyrProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys	100			
Db	241	CAAGACCTGTAGCTGGCGCTCCCGCCAGGTGCGCGCTCATTAACCATGCATTTGC	300			
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg	120			
Db	301	GCCTCTCGGACCTTACCTGGTCACGAGCAAGCGCATGCTTCCTGTGCGCGGACG	360			
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140			
Db	361	GGTATGGCAGGGCAGCGCTGCTTCGCGCGCGCTATCTCTTACTTGAAGGCTCCTCG	420			
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160			
Db	421	GGAGGCGCTCTGCTGTGCGCGCAGGACATGCCGTAGGCATATTACAGCGCGGTATGC	480			
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGluThrThrMet	180			
Db	481	ACCGTGGAGTGGTAAAGCGGTGGACTTCATCCCGTAGAGGCTTAGAGCAACCATG	540			
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200			
Db	541	AGTCCCGGCTGTCTCAGACCACTCTCCCGCCAGCAGTGGCCCGCAGAGTACCAAGTG	600			
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220			
601	GCCACCTGCTGTCCTCCACCGGCGAGCGGTAAAGACCAACCAAGGTCCCGCGCATACGCA	660				
221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240				
661	GCTCAGGCTACAAGGTGCTGCTCAACCCCTCGTTCGCAACAATGGGCTTTGGT	720				
241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260				
721	GCTTACATGTCACAGGCCCATGGATTGATCCTAACATCAGGACTGGGTGAGACAAAT	780				
261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280				
781	ACTACTGGCAGCCGATCAGCTATTCCACCTACGCAAGTTCCTTGGCCGAGCGGGTGT	840				
281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300				
841	TCAGGGGTGCTTATGACATTAATAATTGTGACAGTGCCACTCCACGGATGCAACATCC	900				
301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320				
901	ATCTTGGGCATTGGCACTGCTCTTGACCAAGCAGACCGCGGGCGAGACTGACTGTG	960				
321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal	340				
961	CTCGCCACCGCTACCCCTCGGGCTCCGTCACCTGTCCTCCATCTTAACATCGAGAGGT	1020				
341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360				
1021	GCTCTGTCCACTACGGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGCAAT	1080				
361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380				
1081	AAGGGGGGAGACATCTCATCTTCTGCACCTCAAGAAGAAGTCCGACGAGCTCGCGCA	1140				
381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400				
1141	AACTGTGCGGTGGGGCTCATCGCGGTACTACCGCGGCTTGATGTGCTCGTC	1200				
401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420				
1201	ATCCCGACAGTGTGCTGCTGTCGCACTGACGCGCTCATGACCGGCTTTTACC	1260				
421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440				
1261	GGCAGCTTCGATTGCGGTATAGACTGCACACGTGTGTCCACGACAGCTGCATTCAGC	1320				
441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460				
1321	CTTGACCTTACCTTCCATTGAGACAATCAGCTTCCCGAGTGTCTCTCCCGTACT	1380				
461	GlnArgArgGlyArgThrGlyArgGlyProGlyIleTyrArgPheValAlaProGly	480				
1381	CAAGTCGGGTAGGACTGGCAGAGGAAGCCAGCATCTACAGATTTGTGGCACCGGG	1440				
481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500				
1441	GAGCGTCTCTTGGCATGTTTGACTGCTGTCTCTCTGCGAGTGTATGACGCGGTGT	1500				
501	AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520				
1501	GCTTGGTATGAGCTTACGCGCGCAGACCATCTTGAATTTGGAGGGCGCTTACGGGCTC	1560				
521	ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu	540				
1561	CCGGAGCTTCCGTGTCACAGACCATCTTGAATTTGGAGGGCGCTTACGGGCTC	1620				
541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560				
1621	ACCCACATAGACGGCCACTTCTTATCCAGACAAGCAGAGTGGGGAACCTTCCCTAT	1680				
561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp	580				

Db 1681 CTGGTAGCGTACCAAGCCACCGGTGTGCGGTAGAGCTCAAGCCCTCCCGCGTGGGAC 1740
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAGTGTGTATCGGTCTCAAGCCACCCCTCATGGGCCACACCTCTGCTA 1800
QY 601 TyrArgLeuGlyAlaValGlnAenGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TATAGACTGGCGCGTCTCCAGAAATGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1860
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 1861 ATGACATGTATGTGCGCTGACCTGGAGTGTGTCAGAGTCAAGTCAAGTCAAGTCAAG 1920
QY 641 ValLeuAlaLeuAlaValCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGCGTGTGTCGCGGTATTCCTATCCACAGGTGCGTGTGCTGTGCTGTGCTGCTG 1980
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 ATTGTCTTGTCCGGAAGCGCGCAATCATACCGACAGGGAAGTCTCTTACCGGAGTTC 2040
QY 681 AspGluMetGluGluCys 686
Db 2041 GATGAATGGAGAGTGC 2058

RESULT 2
AX467113 2061 bp DNA linear PAT 17-JUL-2002
LOCUS AX467113
DEFINITION Sequence 1 from Patent WO0214362.
ACCESSION AX467113
VERSION AX467113.1 GI:21900425
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Sallberg, M.
TITLE A hepatitis c virus non-structural ns3/4a fusion gene
JOURNAL Patent: WO 0214362-A 1 21-FEB-2002;
TRIPEP AB (SE)
FEATURES
Location/Qualifiers
1..2061
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Hepatitis C virus NS3/4A coding region"

ORIGIN

Alignment Scores:
Pred. No.: 1,41e-203 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AX467113 (1-2061)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1 ATGGCGCCATATCAGCGCCATATGCCAGTATGCCAGACAGGAGGCGCTTTGGGATGCATAATCACC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 61 AGCTTGACCGCGCGGACAAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 121 GCCACAGACTTCTGTGGCAACCTGCATTAACAGGGGTGTGTGGACTGTCTTACCATGGAGCC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

Db 181 GGAAACAGGACCAFTTGGTCACTTAAGGGTCTCTGTTATCCAGATGTACACCAATGTGGAC 240
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAAGCTCTGATGGCTGGCGCGCTCCCAAGGTGGCGCTCATTAACACCATGCACCTTGC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 301 GGCTCTCTGGAGACCTTACCTGGTCACGAGCAGCGCCGATGTCATTCCTGTAAGAGCTCTCG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 361 GGTGATGGCAGCGGACGCTGCTTCCCGCGGCTATCTCTTACTTGAAGAGCTCTCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGAGGCGCTCTGCTGTGCTCCCGCAGGACATGCGGTAGGCATATTCAGAGCGCGGTATGC 480
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGTAGAGCTTAGAGACCAACCATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 541 AGGTCCCGCGTGTCTCAGACAACTCTCCACCACGAGTGCCTCCAGAGCTACCAAGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 601 GCCACCTGTGATGCTCCACCGCAGCGGTAAAGACCAAGTCCCGCGCGCATGCAC 660
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 661 GCTCAGGGCTACAAGGTGTGGTGTCTCAACCCCTCGTTCGTCGCAACAATGGGCTTGGT 720
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCCAAGGCCCATGGGATTGATCTCAACATCAGGACTGGGTGGTGGACAAAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 781 ACTACTGGCAGCCCGATCAGGTATTCACCTTACCGCAAGTTCCTTCCGACGGCGGTGT 840
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 841 TCAGGGGTGTCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTTGGGCAATTGGCACTGTCTTGAACAAGCAGAGACCGCGGGGGGAGACTGACTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 961 CTCGCCACCGGTACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGAGGTT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGTATTCCTCCCTTGAAGCAAT 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 1081 AAGGGGGGAGACATCTCATCTTCTGCCATCTCAAGAGAGAGTGGCAGAGCTCGCCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 1141 AAACCTGGTTCGGTGGGCGTCAATGCGGTGGCTTACTTACCGCGGCTTGTGTGTCTCGTC 1200
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCGACCAAGTGTGTGCTGTGCTGCGCAACTGACGCGCTCATGACCGGCTTTACC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1261 GCGGACTTCGATTCGGTGATAGACTGCAACACGCTGTGTCAACCCAGACAGTCACTTCAGC 1320

Qy 441 LeuAspProThrPheThrIleGluThrLeuProGlnAspAlaValSerArgThr 460
 Db 1321 CTTGACCTACCTTACCAATGAGCAATCACGCTTCCCGCAGGATGCTGTCTCCGCTACT 1380
 Qy 461 GluArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaProGly 480
 Db 1381 CAACGTCGGGTAGGACTGCGAGAGGAGCCAGGCACTTACAGATTGTGGCCACCGGG 1440
 Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 1441 GAGCGTCTTCTGGCATGTTTGACTGCTGTCTCTGCGAGTGCTATGACGCGGTTGT 1500
 Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAenThr 520
 Db 1501 GCTTGGTATGAGCTTACGCCCGCGAGACCACAGTTAGGCTAGGACATACATGAACACC 1560
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 1561 CCGGCACTTCCCGTGTCCCAAGACCCTTGAATTTTGGAGGGCGTCTTACGGGTCTC 1620
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 1621 ACCCACATAGACGCCCACTTCCATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 1681 CTGTAGCGTACCAAGCACCGTGTGGCTAGAGTTCAGCCCTCCCGCTCGTGGGAC 1740
 Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGAAGTGTTCATCCGCTCTCAAGCCACCCCTCCATGGGCAACACCTTGCTA 1800
 Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 1801 TATAGACTGGCGGTGTTCAGATGAAGTCAACCTGACGCCACCCAGTACCAAGTATATC 1860
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 1861 ATGACATGTATGTCGGTGTACCTGGAGTGTGTCAAGAGTACCTGGTGTGTGGCGGC 1920
 Qy 641 ValLeuAlaLeuAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 1921 GTTCTGGCTCTTGGCCGGTATTCCTATCCACAGGCTGCTGTGTATAGTAGTAGG 1980
 Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 1981 ATTGTCTTGTCCGAAAGCCGCAATCATATCCGACAGGGAAGTCTCTTACCGGGAGTTC 2040
 Qy 681 AspGluMetGluGluCys 686
 Db 2041 GATGAATGGAAGATGC 2058
 RESULT 3
 LOCUS I06434 5360 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 487 from Patent EP 0318216.
 ACCESSION I06434
 VERSION I06434.1 GI:590311
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 5360)
 AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
 TITLE Nanbv diagnostics and vaccines
 JOURNAL Patent: EP 0318216-A1 48 31-MAY-1989
 FEATURES
 source 1..5360
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN

Alignment Scores: 1.74e-200 Length: 5360
 Pred. No.: 3574.00 Matches: 672
 Score: 3574.00 Conservative: 10
 Percent Similarity: 99.42% Mismatches: 4
 Best Local Similarity: 97.96% Indels: 0
 Query Match: 98.78% Gaps: 0
 DB: 6
 US-09-930-591-2 (1-686) x I06434 (1-5360)
 Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 930 CTGCGCGCCATCAGCGCGTACGCCAGCAGACAAGGGCCCTCTTAGGTGCATAATCACC 989
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
 Db 980 AGCCTAACTGCGCGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTCAACTGCT 1049
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 1050 GCCCAACCTTCTTGGCAACGTCATCAATGGGTGTGTGGACTGTCTACACCGGGCC 1109
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1110 GGAACGAGGACCATCGCGTCACCCAGGGTCTCTCATCCAGATGTATACCAATGTAGAC 1169
 Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1170 CAAGACCTTGTGGCTGGCGCGCTCCGCAAGGTAGCGCTCAATGACACCTCGCACTTC 1229
 Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 Db 1230 GGCTCTCGGACCTTACCTGGTCAAGGACGAGCCGATGTCATTCCTGCGCGCGCGCG 1289
 Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 1290 GGTGATAGCAGGCGACCTGCTGTCCCGCGCGGACGCGGTGGGCATATTAGGCGCGCGGTGC 1349
 Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 1350 GGGGTCTCGCTGTGTGCGCGCGGACGCGGTGGGCATATTAGGCGCGCGGTGC 1409
 Qy 151 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 1410 ACCCGTGGAGTGGTAAAGCGGTGACTTTATCCCTGTGGAGACCTAGAGACCAACATG 1469
 Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 1470 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGGCCCAAGAGCTTCCAGGTG 1529
 Qy 201 AlaHisIleHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 1530 GCTCACTCCATGCTCCACAGGCGGCAAGAACCAAGGTCCCGGTGCTATATGCA 1589
 Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 1590 GCTCAGGCTATAGGTGTAGTACTCAACCCCTCTCTGTCTGTCACACACTGGGCTTTGGT 1649
 Qy 241 AlaTyrMetSerLysAlaHisGlyLysAspProAsnIleArgThrGlyValArgThrIle 260
 Db 1650 GCTTACATGTCCAAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAT 1709
 Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 1710 ACCACTGGACCCCATCATCGTACTCCACCTACCGCAAGTTCCTTCCGACGCGGGTGC 1769
 Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 1770 TCGGGGGCGCTTATGACATAATAATTGTGACAGTGCACCTCCAGGATGCCATCC 1829
 Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 1830 ATCTTGGCATCGGCACCTGTCCTTACCAAGACAGACTGCGGGGGGAGACTGGTGTG 1889

QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluVal 340
DB 1890 CTGCCACCCACCCCTCCGGCTCCGTCACTGCCCCATCCCAACATCGAGGAGTT 1949
QY 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle 360
DB 1950 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAGTAATC 2009
QY 361 LysGlyGlyArgHisIlePheCysHisSerIleLysCysAspGluLeuAlaIle 380
DB 2010 AAGGGGGAGAGATCTTCATCTTCTGTCAATCAAGAAGAGTGCAGCACTCGCCGA 2069
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaIleThrArgGlyLeuAspValSerVal 400
DB 2070 AAGCTGTCTGCATTTGGCATCANTCGGTGGCTACTACCGCGTCTTGACGTGTCCGTC 2129
QY 401 IleProThrSerGlyAspValValValAlaIleThrAlaAspAlaLeuMetThrGlyPheThr 420
DB 2130 ATCCCGACCAAGCGGATGTTGTCGTGTGGCAACCGATGCCCTCATGCCGCTATACC 2189
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 2190 GGGCATTCGACTCGGTGATAGTACATACGTGTGTACCCAGACAGTCGATTTACG 2249
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 2250 CTTGACCTTACCTTACCATTTAGACAAATCACGCTCCCGCAGATGCTGTCTCCGCACT 2309
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaProGly 480
DB 2310 CAACGTTCGGGGAGAGTGGCAGGGGAGCCAGGCACTTACAGATTTGTGGACCGGG 2369
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 2370 GAGCGCCCTCCGGCATGTTCCGACTCGTCCGCTCTGTGAGTGTATGACGAGGCTGT 2429
QY 501 AlaTyrThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 2430 GCTTGTGTATGAGTTCACGCCGCCGAGACTTACAGTTAGCTACGACGCTACATGAACACC 2489
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
DB 2490 CCGGGCTTCCCGTGTGCCAGACCATCTTGAAATTTGGAGGGCTCTTTACAGGCTC 2549
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 2550 ACTCATATAGATCCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 2609
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
DB 2610 CTGGTATGATACCAAGCCACCGTGTGCGGTAGGCTCAAGCCCTCCCCCATCGTGGAGC 2669
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 2670 CAGATGTGAAGTGTGTTGTTGCTCCCAAGCCACCTCCATGGGCCAACACCCCTGCTA 2729
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 2730 TACAGACTGGGCGCTGTTTCAAGATGAATCAACCTGACGACCCCACTACCAAAATACATC 2789
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
DB 2790 ATGACATGATGTCGGCCGACCTGGAGTGTGTACGAGACCTGTGGTGTCTGTGGCGC 2849
QY 641 ValLeuAlaLeuAlaIleThrCysLeuSerThrGlyCysValIleValGlyArg 660
DB 2850 GTCTGTGGCTGTGTGGCGGTATTCCTGTCAACAGGCTGTGTGTGTGTGTGTGTGTGTGT 2909
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 2910 GTCTGTGTGGGAGCGGCAATCATCTGACAGGGAAGTCTCTTACCGAGATTC 2969

QY 681 AspGluMetGluGluCys 686
DB 2970 GATGAGATGGAAGAGTGC 2987
RESULT 4
LOCUS I09328 5360 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 8 from Patent WO 8904669.
ACCESSION I09328
VERSION I09328.1 GI:587963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5360)
Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;
FEATURES
Location/Qualifiers
1..5360
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.74e-200 Length: 5360
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0
US-09-930-591-2 (1-686) x I09328 (1-5360)
QY 1 MetaAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 930 CTGGCGCCCATCACGGCTACGCCAGCAGACAGAGGGCCCTCTAGGGTGTATATCACC 989
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 990 AGCCTAATGTCGCGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 1049
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
DB 1050 GCCCAACCTTCTGGCAACGTGATCAATGGGTGTGTGTGACTGTCTACACGGGGCC 1109
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 1110 GGAACGAGGACCATCGGCTCACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 1169
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 1170 CAAGACCTTGTGGGTGCGCGCTCCCGAAGGTAGCGCTCATTTGACACCCCTGCACTGC 1299
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 1230 GGCTCTCGGACCTTACCTGTGTACGAGGACGCGGATGTCATTCCTCGTGGCGCGG 1289
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 1290 GGTGATAGCAGGGGACGCTGCTGTCGCCCGGCCCATTTCTACTTGAAGAGGCTCCTCG 1349
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 1350 GGGGTGCGCTGTGTGTGCGCGCGGSCAGCGCGTGGGCATATTTAGGCGCGGTGTGC 1409
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 1410 ACCGCTGGAGTGGTAAAGGGGTGGACTTTATCCCTGTGGAGAACCCTAGAGACAACCATG 1469
QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
DB 1470 AGTCCCGGGTGTTCACGGATATCTCTCTCCACCATGTAGTCCCGCCAGAGCTTCCAGGTG 1529

201	Qy	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
1530	Db	GCTCAGCTCATGCTCCACAGCAGCGGCAAAAGCACCAAAGCTCCGGCTGATATGCA	1589
221	Qy	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaAlaThrMetGlyPheGly	240
1590	Db	GCTCAGGGCTAAGAAGTGCTAGTACTCAACCCCTCTGTTGCTGCACACATGGGCTTTGGT	1649
241	Qy	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
1650	Db	GCTTACATGTCCTCAAGGCTCATGCGATCGATCCTAAACATCAGACCGGGGTGAACAACATT	1709
261	Qy	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
1710	Db	ACCACGTGGCAGCCCCATCAGTACTCCACTCAGCAAGTTCTCTCCGACGCGCGGTGC	1769
281	Qy	SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
1770	Db	TCGGGGGGCGCTTATGACATAATAATTGTGACGAGTGCCTCCACGGATGCACATCC	1829
301	Qy	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgIleuThrVal	320
1830	Db	ATCTTGGGCATCGCACTGCTCTGTACCAAGCAGACAGACTCGGGGGCGAGACTGGTTGTG	1889
321	Qy	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
1890	Db	CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGTT	1949
341	Qy	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
1950	Db	GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGCTATCCCCCTCGAAGTAATC	2009
361	Qy	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla	380
2010	Db	AAGGGGGGAGACATCTCATCTTCTGTCANTCAAGAGAGAACTGGCAGCACTCGCCGCA	2069
381	Qy	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
2070	Db	AAGCTGGTCGANTGGGCATCAATGCCGTGGCCTACTACCCGGGTCTTGACGTGTCGTC	2129
401	Qy	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
2130	Db	ATCCGACCAAGCGCGATGTGTGCTCGTGGCAACCGATGCCCTCATGACCGGCTATACC	2189
421	Qy	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
2190	Db	GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCAGACAGTCGATTTCCAG	2249
441	Qy	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
2250	Db	CTTGACCCCTACCTTCACCATTTGACAAATCACGCTCCCCAGAGATGCTGCTCCCGCACT	2309
461	Qy	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
2310	Db	CAACGTCGGCGAGGACTGGCAGGGGAGCCAGGCATCTACAGATTGTGGCACCGGGG	2369
481	Qy	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
2370	Db	GAGCGCCCTCCGGCATGTCGATCTCGTCCGTCTCTGTGAGTGCTATGACGAGAGCTGT	2429
501	Qy	AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
2430	Db	GCTTTGGTATGAGCTACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC	2489
521	Qy	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
2490	Db	CCGGGGCTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGCTCTTTTACAGGCCTC	2549
541	Qy	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
2550	Db	ACTCATATAGATGCCCACTTCTATCCAGACAAGCAGAGTGGGAGAACCTTCTCTTAC	2609
561	Qy	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp	580

Db	2610	CTGGTAGCTACCAAGCCACGGTGGCTAGGGCTCAAGCCCTCCCATCGTGGAC	2669
Qy	581	GlnMetTrpLysCysLeuIleAArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	2670	CAGATGTGGAAGTGTGTGATTTCGCCTCAAGCCACCCCTCCATGGGCAACACCCCTGCTA	2729
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	2730	TACAGACTGGGCGCTGTTTCAGATGAATACACCTGACGACCCAGTCCACCAATACATC	2789
Qy	621	MetThrCysMetSerXlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	2790	ATGACATGATGTCGGCGGACCTTGGAGGTGCTCACCAGACACCTGGGTGCTGTTGGCGGC	2849
Qy	641	ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	2850	GTCTTGCTCTTTGGCGCGGTATTGCTGTCAACAGGCTGCTGTGTCATAGTGGGAGG	2909
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	2910	GTCGTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGAAGTCTCTTACCAGAGATT	2969
Qy	681	AspGluMetGluGluCys 686	
Db	2970	GATGAGATGGAGAGTGC 2987	
RESULT 5			
LOCUS	I06440	6785 bp DNA linear	PAT 02-DEC-1994
DEFINITION	Sequence 54 from Patent EP 0318216.		
ACCESSION	I06440		
VERSION	I06440.1	GI:590312	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 6785)		
AUTHORS	Houghton, M., Choo, Q.-L. and Kuo, G.		
TITLE	Nanb diagnostics and vaccines		
JOURNAL	Patent: EP 0318216-Al 54 31-MAY-1989;		
FEATURES	Location/Qualifiers		
source	1..6785		
ORIGIN	/organism="unknown"		
Alignment Scores:			
Pred. No.:	2,32e-200	Length:	6785
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	6	Gaps:	0
US-09-930-591-2 (1-686) x I06440 (1-6785)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	1203	CTGGCGCCCATCACGGCGTACGCCACACAGGGGCTCTCTAGGGTGCAATATCACC	1262
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerAla	40
Db	1263	AGCCTAACTCGCGGGGCAAAAACCAAGTCGAGGGTGAGGTCAGATTGTGTCAACTGCT	1322
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
Db	1323	GCCCAAACTCTCTGGCAACGTGCATCAATGGGGTGCTCGAGTCTTACCACGGGCCC	1382
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	1383	GGAACGAGGACCATCGCTCACCACGGGTCCTGTCAACAGATGTATACCAATGTAGAC	1442

Qy 81 GlnAspLeuValGlyThrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1443 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGAGTAGCGCTCATTCACACCCCTGCATTGC 1502
 Qy 101 GlySerSerAspLeuValLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 Db 1503 GGCTCCTCGACCTTACCTTGGTTCACGAGGACGCGCATGCTATTCCGTCGCGCGCGG 1562
 Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuIleGlySerSer 140
 Db 1563 GGTGATAGCAGGCGCAGCTGCTGTCGCGCGCGCCCATTTCTTACCTTGAAGGCTCCTCG 1622
 Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 1623 GGGGTCCTGTTGTGTCGCGCGCGCGCAGCTGTCGCGCATATTTAGGCGCGCGGTGTC 1682
 Qy 161 ThrArgGlyValAlaIleAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 1683 ACCGCTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACCATG 1742
 Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 1743 AGTCCCGCGGTGTTCAGGATAACTCCTCTCCACCAAGTAGTGCCCGCAGAGCTTCCAGGTG 1802
 Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLeuSerThrIleValProAlaAlaValAla 220
 Db 1803 GCTCACCTCCATGCTCCACAGCAGCGGCAAGCACCAAGCTCCGCGCTGCATATGCA 1862
 Qy 221 AlaGlnGlyTyrIleValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 1863 GCTCAGGGCTATTAAGGTGTGTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 1922
 Qy 241 AlaTyrMetSerIleAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 1923 GCTTACATGTCAGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAT 1982
 Qy 261 ThrThrGlySerProIleThrThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 Db 1983 ACCACTGGCAGCCCATCATCGTACTCCACTACGGCAAGTTCCTTCCGACGCGCGGTGC 2042
 Qy 281 SerGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 2043 TCGGGGGCGGCTTATGACATAATAATTGTGACGAGTGCCACATCCACGATGCCATCC 2102
 Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 2103 ATCTTGGCATCGCATGTCCTTGTACCAAGCAGAGACTCGGGGCGAGAGCTGGTTGTG 2162
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 2163 CTCGCCACCGCCCATCCGCTCGGCTCCGTCACCTGTCGCCCATCCCAACATCGAGGAGTT 2222
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2223 GCTCTGTCCACACCGGAGATCCCTTTTACGGCAAGGCTATCCCTCCGATGAATC 2282
 Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAla 380
 Db 2283 AAGGGGGGAGACATCTCATCTCTCTCATTTCAAGAAGATGCGACCACTCGCGCGCA 2342
 Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 2343 AAGCTGTGTCATTTGGCATCAATGCGGCTTACTACCGCGGTCTTACGCTGTCGCGTC 2402
 Qy 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2403 ATCCCGACGACGCGGATGTTGTCGTGTCGCAACCGATGCCCTCATGCCGCTATACC 2462
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 2463 GCGCATTCGACTCGGTGTAGACTGCAATACGTGTGTACCCAGACAGCTCGATTCACG 2522
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460

Db 2523 CTTGACCCCTACCTTCAACATTGAGACAATCACTCCGCTCCCGCAGGATGCTGTCTCCCGCACT 2582
 Qy 461 GlnArgArgGlyValThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 2583 CAACGTCGGGCGAGACATGCGAGGGGAGCGAGCACTCTACGATTTGTGGCACCGGGG 2642
 Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 2643 GAGCGCCCTCCGCGCATGTCGACTCGTCCGCTCTCTGTGAGTGTATGACGCGAGCTGT 2702
 Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 2703 GCTTGGTATGAGCTCACCCCGCAGACTACAGTTAGCTACGAGCGGTATCATGAACACC 2762
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
 Db 2763 CGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGGCTCTTTACAGGCCTC 2822
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyGluAsnLeuProTyr 560
 Db 2823 ACTCATATAGATGCCACTTTCTATCCAGACAAGCAGAGTGGGAGAACCTTCCCTTAC 2882
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
 Db 2883 CTGGTAGCTACCAAGCCACCGTGTGCTAGGGCTCAGGCCCTCCCCCATCGTGGGAC 2942
 Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrIleHisGlyProThrProLeuLeu 600
 Db 2943 CAGATGTGAAAGTGTGTTGATTGCGCTCAAGGCCACCTCCATGGGCCCAACCCCTGCTA 3002
 Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysThrIle 620
 Db 3003 TACAGACTGGCGCTGTTTCAAGATGAATACCTTACGACCGCACCCAGTCCCAAAATACATC 3062
 Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrThrThrValLeuValGlyGly 640
 Db 3063 ATGACATGCATGTCCGCGACCTGGAGGTGCTCAGGACCTCGGTGCTCGTGGGCGC 3122
 Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrThrGlyCysValValIleValGlyArg 660
 Db 3123 GTCTTGGCTGCTTTGGCGCGCTATTGCTGTCAACAGGCTCGGTGCTCATAGTGGGCGG 3182
 Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 3183 GTCGCTTTGTTCGGGAAGCGCGCAATCATCTCAGGGAAGTCCCTTACCGAGAGTTC 3242
 Qy 681 AspGluMetGluGluCys 686
 Db 3243 CATGAGATGGAAGATGC 3260

RESULT 6
 I09329
 LOCUS I09329 6785 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence (10) from Patent WO 8904669.
 ACCESSION I09329
 VERSION I09329.1 GI:587964
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1. (Bases 1 to 6785)
 AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
 JOURNAL Patent: WO 8904669-A 10 01-JUN-1989;
 FEATURES
 source
 1. 6785
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 2.32e-200 Length: 6785
 Score: 3574.00 Matches: 672

Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	6	Gaps:	0
US-09-930-591-2 (1-686) x 109329 (1-6785)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	1203	CTGCGCGCCCATCAGCGGTACGCCAGCAGACAGAGGGGCTCTAGGGTGCAATACCC	1262
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGlyGluValGlnIleValSerThrAla	40
Db	1263	AGCCTAACTGCGCGGCAAAACCAAGTGGAGGGTGGAGTCCAGATTGTCTCAACTGCT	1322
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysIleThrValTyrHisGlyAla	60
Db	1323	GCCCAAACTTCCTGGCAAGTGCATCAATGGGGTGTGTGGACTGTCTACCAACGGGGC	1382
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	1383	GGAAACGAGGACCATCGGCTCACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC	1442
Qy	81	GlnAspLeuValGlyTyrProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	1443	CAAGACCTTGTGGCTGGCGCTCCCAAGGTAGCGCTCATGTGACACCTGCATTTGC	1502
Qy	101	GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
Db	1503	GGCTCCTCGGACCTTACCTGGTCACGAGCAGCGCGATGCTATCCCGTGGCGCGCG	1562
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	1563	GGTGATAGAGGGGACCGCTGCTGTCGCCCGCGCCCAATTCCTACTGAAAGCTCTCTG	1622
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	1623	GGGGTCCGCTGTGTGCCCCGGGGCACCGCTGGGCATATTTAGGGCGCGGTGTGC	1682
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	1683	ACCCGTGGAGTGGTAAAGCGGTGGACTTATCCCTGTGGAGAACCTTAGAGACCAACATG	1742
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	1743	AGTCCCGGCTTTCAGGGTAATCTCTCCACAGTAGTGGCCCCAGAGCTTCCAGGIG	1802
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	1803	GCTCACCTCCATGCTCCACAGGCGAGCGGCAAAAGCACCAAGTCCCGGCTGCATATGCA	1862
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	1863	GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTCTGCAACACTGGGCTTTGT	1922
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	1923	GCTTACATGTCCTCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT	1982
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys	280
Db	1983	ACCACTGGAGCCCATACAGTACTCCACCTACGGCAAGTTCCTTGGCGAGCGGGGTGC	2042
Qy	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	2043	TCGGGGGGCGCTTATGACATAATAATTTGTGACAGTGGCCACTCCACGGATGCCATCC	2102
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyLeuThrVal	320
Db	2103	ATCTTGGGCATCGGCATGCTCTTGACCAAGAGAGACTCGGGGGGAGACTGGTTGTG	2162
Qy	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal	340
Db			
Db	2163	CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCCATCCCAACATCGAGGAGTT	2222
Qy	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	2223	GCTCTGTCCACCCAGGAGAGATCCCTTTTACGCAAGGCTATCCCTCGAAGTAATC	2282
Qy	361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
Db	2283	AAGGGGGCAGACATCTCATCTTCTGTCAATCAAGAGAGAGTCCGACCACTCGCCGCA	2342
Qy	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	2343	ANGCTGTCTCGATGGGCATCAATCGGTGGCTTACTACCGCGTCTTGAGTGTCCGTC	2402
Qy	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
Db	2403	ATCCGACACGCGCGATGTTGTGCTGGCAACCGATGCCCTCATGACGGGTATACC	2462
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	2463	GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCTGATTCAGC	2522
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	2523	CTTGACCTACCTTCAACATTGAGCAATCAGCTCCCGAGGATGCTGTCTCCCGCACT	2582
Qy	461	GlnArgArgGlyArgThrGlyArgGlyArgGlyIleTyrArgPheValAlaProGly	480
Db	2583	CAACGTGGGGCAGGACTGGCAGGGGGAAGCCAGGATCTACAGATTTGTGGCACCGGG	2642
Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	2643	GAGCGCCCTCCGCACTGTTGCTGCTCTCTGCTGCTGCTATGACGAGGCTGT	2702
Qy	501	AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	2703	GCTTGGTATGACTCACGCCCGCGAGACTACAGTATAGGCTACAGGCTACATGAACACC	2762
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu	540
Db	2763	CGGGGCTTCCGCTGTGCCAGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCTC	2822
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	2823	ACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC	2882
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp	580
Db	2883	CTGGTAGGCTACCAAGCCCGCTGTGCTAGGCTCAAGCCCTCCGCCATCGTGGGAC	2942
Qy	581	GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	2943	CAGATGTGGAAGTGTGTTGATTGCTCAAGCCCACTCCATGGGCCAACACCCCTGCTA	3002
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	3003	TACAGACTGGGCGCTGTTCAGATGAATACCTCCCTGACGACCCAGTCCCAATATCATC	3062
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrThrValLeuValGlyGly	640
Db	3063	ATGACATCCATGTCGGCGGACCTGGAGTCTGCACGACACCTGGGTGCTGTGGGGC	3122
Qy	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	3123	GTCTCTGGCTGTTTGGCGCGTATTGCTGTCAACAGGCTCGGTGGTTCATAGTGGGAGG	3182
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	3183	GTCTGTCTTGTCCGGGAGCCCGCAATCATCTCACAGGGAAGTCTCTACCGAGAGTTC	3242
Qy	681	AspGluMetGluGluCys	686
Db	3243	GATGAGATGGAAGAGTGC	3260

2328	GCTCACCTCCATGCTCCACAGGACGGGCAAAAGACCAAGGTCCCGGCTGCATATGCA	2387
221	AlaGlnGlyTyrIysValLeuValLeuAenProSerValAlaAlaThrMetGlyPheGly	240
2388	GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTGGT	2447
241	AlaTyrMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
2448	GCTTACATGCTCAAAGGCTCATGGATCGATCTTAAACATCAGACCGGGTGAGAACAT	2507
261	ThrThrGlySerProIleThrTyrSerThrTyrGlyIlePheLeuAlaAspGlyCys	280
2508	ACCACTGGAGCCCATCAGTACTCCACCTACGCGCAAGTTCCTTCCGACGCGGGTGC	2567
281	SerGlyGlyAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer	300
2568	TCGGGGGGCGCTTATGACATAAATAATTGTGACAGTGCACCTCCACGATGCCACATCC	2627
301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
2628	ATCTGGGCATCGGCACTGTCTTGACCAAGCAGAGACTGGGGGGGAGACTGGTTGTG	2687
321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
2688	CTCGCCACCGCACCCCTCGGGGCTCGTCACTGTGCCCATCCCAACATCGAGAGGTT	2747
341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyIleAlaIleProLeuGluAlaIle	360
2748	GCTCTGTCACCCAGGAGATCCCTTTTACGGCAGGCTATCCCTCCAGATATC	2807
361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAla	380
2808	AAGGGGGGAGACATCTCATCTCTGTCATTAAGAGAGAGTGCAGCACTCGCGCA	2867
381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
2868	AAGCTGGTCGATTGGGCATCAATGCCGTGGCTTACTACCGGGTCTTGACGTGTCGTC	2927
401	IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr	420
2928	ATCCGACAGCGGAGATGTTGTCGTGGCAACCGATGCCCTCATGACCGGCTATACC	2987
421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
2988	GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACAGACAGTCGATTTT	3047
441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
3048	CTTGACCTTACCTTACCATTTGAGACAAATCACGCTCCCCCAGAGATGCTGTCTCCG	3107
461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
3108	CAACGTGGGGGACGACTGGCAGGGGAGACCGAGCATCTACAGATTTGTGGCACCGGG	3167
481	GluArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTyrAspAlaGlyCys	500
3168	GAGCGCCCTCCGGCATGTTGCACTCGCTCGCTCTGTGAGTGCTATGACGAGGCTGT	3227
501	AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
3228	GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGCTACGAGCGTACATGAACACC	3287
521	ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu	540
3288	CCGGGGCTTCCCGTGTCCAGGACCATCTTGAATTTTGGAGGGGCGCTCTTTACAGCC	3347
541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
3348	ACTCATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC	3407
561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp	580
3408	CTGGTAGCTACCAAGACCGCTGTGGCTAGGGGCAAGCCCTCCCGCATCGTGGGAC	3467

Qy 581 GlnMetTrpLysCysLeuIleAlaValGlnAsnGluValThrLeuThrHisProValThrHisGlyProThrProLeuLeu 600
 Db 3468 CAGATGTGGAAGTGTGTTGATTGGCTCAAGCCACACCCCTCCATGGCGCAACACCCCTGCTA 3527
 Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrHisGlyProThrProLeuLeu 620
 Db 3528 TACAGATGGCGGTGTTCAGAAATGAAATCACCTGACGACCCAGTCACCAATACATC 3587
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 3588 ATGACATGATGTCGGCGGACCTGGAGTGTGTCACGAGACACCTGGTGTGTCGTTGGCGGC 3647
 Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValLeuValGlyArg 660
 Db 3648 GTCTGTGCTGCTTTGGCGCGGTATGCTGTCTACACAGCGTCGCTGTGTCATAGTGGGCAGG 3707
 Qy 661 IleValLeuSerGlyLysProAlaIleProAlaIleProAspArgGluValLeuValTrpArgGluPhe 680
 Db 3708 GTCTGTGCTGCTTTGGCGGACCTGGAGTGTGTCACGAGACACCTGGTGTGTCGTTGGCGGC 3767
 Qy 681 AspGluMetGluGluCys 686
 Db 3768 GATGAGATGGAAGATGC 3785
 RESULT 8
 I09331
 LOCUS I09331 7310 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 15 from Patent WO 8904669.
 ACCESSION I09331
 VERSION I09331.1 GI:587966
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 7310)
 AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
 JOURNAL Patent: WO 8904669-A 15 01-JUN-1989;
 FEATURES
 source
 1. 7310
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,54e-200 Length: 7310
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x I09331 (1-7310)
 Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 1728 CTGGCGGCCCATCAGCGGTACGCCAGACAGGAGGCGCTCTAGGGTGCAATACCC 1787
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
 Db 1788 AGCCTAACTGGCGGGACAAACCAAGTGGAGGTGAGGTGTCAGATTGTGTCACCTGCT 1847
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThrHisGlyAla 60
 Db 1848 GCCCAACCTTCTGGCAACGTGCATCAATGGGGTGTGCTGATCCAGATGATACCATGTAGAC 1907
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1908 GGAACGAGGACCATCGGTACCAAGGGTCTCTGATCCAGATGATACCATGTAGAC 1967
 Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1968 CAAGACCTTGTGGGCTGGCGGCTCCGCAAGGTAGCGGCTCATTGACACCCCTGCACTTGC 2027

Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAlaArgArg 120
 Db 2028 GGCCTCTCGGACCTTTACCTGGTCACGAGGACCCGATGTCAATTCGCGTCGCGCGCGG 2087
 Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 2088 GGTGATACAGCGGGCAGCTCTGTGCGCCCGGCCCAATTTCTACTTGAAGGCTCTCTCG 2147
 Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 2148 GGGGGTCCGCTGTTGTGCGCGCGGACGCCGTGGCATATTTAGGCGCGCGGTGTGC 2207
 Qy 161 ThrArgGlyValAlaAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
 Db 2208 ACCCGTGGAGTGGCTAAGCGGTGACTTTATCCTGTGGAGAACCTTAGAGACAACCATG 2267
 Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 2268 AGGTCCCGCGTGTTCACGGATAACTCTCTCCACAGTAGTCCCGGAGCTTCCAGGTG 2327
 Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 2328 GCTCACCTCCATGCTCCACAGGCGGCAAGACCAAGGTCCTCGGCTGCAATATGCA 2387
 Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 2388 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTGGT 2447
 Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 2448 GCTTACATGTCCAAAGGCTCATGGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAA 2507
 Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 2508 ACCACTGGCAGCCCCCATCAGTACTCCACCTACGCGAAGTCTTTCGCGCGCGGGTGC 2567
 Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 2568 TCGGGGGCGGCTTATGACATAATAATTTGTGACAGTGCCACTCCAGGATGCCACATCC 2627
 Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 2628 ATCTTGGCATCGGCATGCTCTTGACCAAGACAGACTCGCGGGGCGAGACTGTTGTG 2687
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 2688 CTGCGCCACCGCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGAGGTT 2747
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2748 GCTCTGTCCACCGGAGAGATCCCTTTTACGCGAAGGCTATCCCCCTCGAAGTAAATC 2807
 Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 2808 AAGGGGGGAGACATCTCATCTTCATCAAGAGAGAGTCCGACGAACTCCCGCGCA 2867
 Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 2868 AAGTGTGTGATGGGCATCAATCGCGTGGCTTACCGCGGTCTTGACGTGTCGCTC 2927
 Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2928 ATCCGACGACGCGGATGTTGCTGTGGCAACCATGCCCTCATGACCGGTATACC 2987
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 2988 GCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGCTGATTTTCCAG 3047
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 3048 CTTGACCTCACTTCACTTACCAATTGAGACAAATCAGCTCCCCCAGGATGCTGTCTCCGCACT 3107

```
Qy 461 GlnArgGlyArgThrGlyArgGlyArgProGlyLeuValAlaProGly 480
Db 3108 CAACGTCGGGCGAGGACTGCGACGGGGAAGCCAGGCATCTACAGATTGTGGCACC GGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATGTTGCACTCGTCGTCCTCTGTGAGTGCTATGACGAGGCTGT 3227
Qy 501 AlaTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAenThr 520
Db 3228 GCTTGGTATGAGTCAAGCCCGCGAGACTACAGTTAGGTCAGAGCGTACATGAACACC 3287
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTCCCGTGTGCCAGGACCATCTTGATTTTGGAGGGCGCTCTTACAGGCCCTC 3347
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAenLeuProTyr 560
Db 3348 ACTCATATAGATGCCACTTCTATCCACAAAGCAGAGTGGGAGAACCTTCCTTAC 3407
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 3408 CTGGTAGCGTACCAGCCCGTGTGGCTAGGGCTCAAGCCCTCCCCATCGTGGGAC 3467
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCCACTCCATGGGCCAACACCCCTGTCTA 3527
Qy 601 TyrArgLeuGlyAlaValGlnAenGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTGGCGCGCTGTTCAGATGAATCAACCTCAGCCACGACCCATCAACAAATCATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 3588 ATGACATGATGTCGGCGGACCTGGAGTGTCTACGAGCACCTGGTGTGGTGGCGGC 3647
Qy 641 ValLeuAlaLeuAlaTyrCysLeuSerThrGlyCysValValleValGlyArg 660
Db 3648 GTCTGCTGCTTTGGCGCGTATTGCTCTCAACAGGCTGCTGATAGTGGCGAGG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTGCTGTTCGGGAAGCGGCAATACCTACGAGGGAAGTCTCTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGlyCys 686
Db 3768 GATGAGTGAAGAGTGC 3785

RESULT 9
HPCPOLYP
LOCUS
DEFINITION
Hepatitis C virus polyprotein gene, partial cds.
ACCESSION
M32084
VERSION
M32084.1
GI:329875
KEYWORDS
polyprotein
Hepatitis C virus
ORGANISM
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1 (bases 1 to 7310)
AUTHORS
Choo, Q.-L., Richman, K. and Han, J.
TITLE
The nucleotide sequence of the Hepatitis C viral genome
JOURNAL
Unpublished (1990)
COMMENT
Original source text: Hepatitis C virus, cDNA to viral RNA, clones
K9-1 through 15e, isolated from chimpanzee (individual 910) blood
plasma.
Draft entry and printed sequence for [1] kindly submitted by
M'Houghton, 22-FEB-1990; Chiron Corporation, 4560 Horton Street,
Emeryville, CA 94608.
FEATURES
Location/Qualifiers
1..7310
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
```

CDS

```
/db_xref="taxon:11103"
<1..>7310
/note="polyprotein"
/codon_start=3
/protein_id="AAA45677.1"
/db_xref="GI:329876"
/translat="GCPERLASCRPLTDFDQMGPISVANGSGDQDPYQWHPKPC
GLVPAKSCGPEVYCFPSVVGTTDRSGAFYSWGENDTDFVNNRTRPLGNWFGC
TWMSNTGFTKVGCGAPPVGGAGNNTLHCPDCKRHPDTRATSCGGSGPWITPCLVD
YFRLMHTPCNTINTYIKIRMYGVGEHLEACNCTRGERCDLDRSLSLLLT
TTOQVLPCSTFTLSTGLIHQNIIVDQYLVGVSSIASWAIKWEYVLLVLL
ADAVSCLMWMLLIISOAALENLVINAASLAGTGLVSLVFFCFVAVLKGKWP
GAVTYGMPPLLLALLPORVALDEVAASCGVVLGLMALTLSPYKRYISWC
LWNTQYFLTYVEAOLHWIPLNVRGGRDAVILAMCAVHTPLVDITKLILAVGPELW
ILOSLLVKVPYFVRVQGLLRFCALARKMIGHYVQWVILKLGALTGYVYNHILPLRD
WAHNGRLDAVAVPVSQMETKLTITWGAADTAACGDIINGLPVSARGRILLUGPAD
GMVSKGRMLAPITAYAQOTRGLGCIITSLTRGDKQVEGEVITFAQTFLATCI
NGVCTWYHVGAGTRTIASPKGPVQMTNVQDLPWPAPOGSRSLTPTCGSDLYL
VTRIAADVPRRRGDSRGLSPRPISYLKSGSGGLICPAHAGVIFRAACVTRGVA
KAVDPIPVENLETTMRSPVFTDSSPPVPOSFOVHLHAPTGSKGKTKVAAVAAQGS
KVLVLAAPSVAATLFGFAYSKAGIDENIETGVTITGSPITYSTYKGLADGQS
CGADILICDECHSDATSLIGITVLDAQETAGARLVLATATPPGSTVTPHNIEE
VALSTTGEPYFKAIPLEVIKGRHLIFCHSKKKDELAALVALGINAVAYRGLD
VSVPTSGDVVVAATDALMTGTDGDFSDVDCNTCTQTVDFSLDPTFTIITLPQD
AVSTQRRTGRGKPGIYFVAPGERPSGDFSSVLCECYDAGCAYELTAPETVAR
LRAYNTPLGPVQCQDHLFMEGVFTGLTHIDAFELSTQKSGENLPLVATQATVCR
ACAPPSWDQMWKGLIRLPTLHGTPLLVLGAVONEITLTHPVTKYIMTCMSADLE
VYTSVLVGGVLAALAAACLSLTCGVIVGSLGKPAIIPDREVLYRBEDEECS
QHLPIYIGQMMLAEQFKALGLOLTSRQREVIAPAVQTNWQKLETFWAKHMNFI
GIQIAGLSLTPGNPAIASLMAFTAAVTSPLTSTQTLNFIILGNGWAAQALAPGAATA
FVAGLAGAATGSGVLGKLDILAGYGAGVALVAFKIMSGEVPTEDLVNLLPAI
LSPGALLVGVVQCAAILRRHVGFGEVQWNNLFIASFGRNHVSPTHYVPSDAAARV
TALSLSLTVQALLRRLHOWISSECTPCSGSLRDIWDWICEVLSDFKTLWAKLMPO
LPGIPFVSCORGKGVMEVDGIMETRCHGAEITGHVKNGTMRIVGPTCRNMSGTF
PINATYTGCTPLPAPNYTFALWASAEYVEIRQVGDGHVYVTGMTDNLKCPQSPS
PEFTFELDKRHLRPAKPLRESEVSRVGLHEYPVSGQPCPEPFDVAVLVSLMT
DPSTHABAGRLIARGSPVSSASSQSLAKATCTANHDSPOALELLEANLWR
QMGENTIRVESNWLIDFSLVAEEDEREISVPAILKSRRAQALVWALRPD
YNPLFETWTKDYEPPVPHGCPPLPPKPPPPKRRKTVLTTESTLSTALALATR
SPPGSSVTGITGDNNTSSPAPSGPDSAEYSVMPLEGGPGDPLSDQSGWSTV
SSEANEDVCCMSYSWTGALVTPCAAEOKLPINALSNLSLRHNLVLTSTSRAC
QKQKVPDLQVLDSHYQDVLKEVKAASKVXANLLSVEACSLTFPHSKSKFVGC
AKDVRCHARKATHINSWKLDDNVTPTDITIMAKNEVFCVQPEKGRPAKLIIV
PDLGVRCENKALYDVVTKLPLAYGSSYGFQSPQRVEPLVQAWSKKTPMGFSYD
TRCFDSTVEDIRTEALYQCDLDPQARVAIKSLTERLYVGGFLNRSNGCYRR
CRASGLTSCGNTLTCVTKARAAACRAAQDCTMLVCGDLYVICSAGVQEDAAAL
RAFTAMTRYAPPPDPOPEYDLELITSCSNVSVADHDGAKRVYLTRODPTPLAR
AAWETARHTPVNSMLNLIIMFAPTLWARMILMTHFFSVLIARDQLEQALDCEIYGACY
SIEPLDLPPIIQL"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,54e-200 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 14 Gaps: 0

US-09-930-591-2 (1-686) x HPCPOLYP (1-7310)
Qy 1 MetaAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 :::: 1728 CTTGGCGCCCATCAGCGGTGATCGCCAGACAGAGGGCCCTCTAGGGTGCATTAATCACC 1787
Qy 21 SerLeuThrGlyArgAspLysAenGlnValGluGlyGluValGlnIleValSerThra 40
Db 1788 AGCCTAACTGGCCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 1847
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAenGlyValCystTyrThrValTyrHisGlyAla 60
Db 1848 GCCCAAAACCTTCCTGGCAACGTGCATCAATCGGGGTGTGCTGGACTGTCTACACGGGGCC 1907
```

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1908 GGRACGAGGACCATCGCGTCAACCAAGGTCCTGTCTCCAGATGATATCAATGATAGAC 1967
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1968 CAAGACCTTGTGGCTGCGCGCTCCGCAAGGTAGCGGCTCATGTACACACCTGCATTCG 2027
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAlaArgAspArg 120
 Db 2028 GGCTTCCTCGACCTTACCTTGTGTGCGGAGGACCGCGCATGTCTTCCTCGCGCGCGG 2087
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 2088 GGTGATAGCAGGGGCGAGCTGCTGTCGCGCGCGGCGCCATTTCTTACTTGAAGGCTCTCTCG 2147
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 2148 GGGGGTCCGCTGTGTGCGCGCGGCGGACCGCGTGGGCATATTTAGGGCGCGGCTGTGC 2207
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 2208 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCTCTGCGAGAACCCTAGAGACAACCATG 2267
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 2268 AGGTCCCGCGGTTCACGGATAACTCTCTCTCCACAGTAGTGGCCCCAGAGCTTCCAGGTG 2327
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 2328 GCTCACCTCCATGTCTCCACAGGCGGCGGCAAGACCAAGGTCCCGGTGCATATGCA 2387
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 2388 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTCTGTCGAACACACTGGGCTTTGGT 2447
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaArgThrIle 260
 Db 2448 GCTTACATGTCAGAGCTCATGGAGTGAUCCTAACATCAGGACCGGGGTGAGACAAATT 2507
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 2508 ACCACTGGCAGCCCATCACGTACTCACCTACGGCAAGTTCCTTGGCGGCGGCGGTGC 2567
 QY 281 SerGlyGlyAlaTyrAspIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 300
 Db 2568 TCGGGGGCGCTTATGACATAATTAATTGTGACGAGTGCCACTCCCGATGCCCATCC 2627
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 2628 ATCTTGGGCATCGGCACCTGCTTGTACCAAGCAGAGACTGCGGGGCGAGACTGGTTGTG 2687
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 2688 CTCGCCACCGCCCTCCCGGCTCGCTCACTGTGCCCCCATCCCAACATCGAGGAGGTT 2747
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2748 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCGCTCGAGTAATC 2807
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 2808 AAGGGGGGAGACATCTCATCTTCTGTCAITCAAAGGAAGAGTGGCAGAACTCGCGCA 2867
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 2868 AAGCTGTGCGATTGGGCGATCAATGCGGCTACTACCGGCTCTTACGCTGTCGTC 2927
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2928 ATCCGACACGCGGCGATGTTGTGCTGTCGTAACCGATGCCCTCATGACCGGCTATACC 2987
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

Db 2988 GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACCCAGACAGTGCATTCAGC 3047
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 3048 CTGTACCTTACCTTCCATTGAGACAATACGCTCCCGCAGGATGCTGTCTCCGCACT 3107
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 3108 CAACGTCGGGGCAGGACTGGCAGGGGGAAGCAGGCACTTACAGATTGTGGCAACCGGG 3167
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 3168 GAGCGCCCTCCGCGCATGTTGACTTCGCTCCGCTCTGTGAGTGTATGAGCGAGGCTGT 3227
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 3228 GTTGTGTATGAGCTCACCGCCCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 3287
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
 Db 3288 CGGGGGCTTCCGCTGTGCGAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCTC 3347
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 3348 ACTCATATAGATGCCACTTTCTATCCAGACAAGCAGAGTGGGAGAACCTTCTTAC 3407
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerThrAsp 580
 Db 3408 CTGTGTAGGCTACCAAGCCACCGTGTGCTAGGCTCAAGCCCTCCCGCATCTGGGAC 3467
 QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 3468 CAGATGTGGAGTGTGTGATTCGCTCAAGCCACCTCCATGGGCCAACACCTCTGCTA 3527
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 3528 TACAGACTGGGCGCTGTTCAGATGAAATCACCTGACGACCCAGCTACCAATACATC 3587
 QY 621 MetThrCysMetSerAlaAspLeuValValThrSerThrThrTrpValLeuValGlyGly 640
 Db 3588 ATGACATGCATGTGCGCGACCTGGAGTGTGTGAGTGTGTGAGTGTGTGTGTGTGTGT 3647
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 3648 GTCTGTGCTGT 3707
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 3708 GTCTGT 3767
 QY 681 AspGluMetGluGluCys 686
 Db 3768 GATGAGATGGAGAGTGC 3785

RESULT 10
 I08294
 LOCUS
 DEFINITION Sequence 1 from Patent-EP 0388232
 ACCESSION I08294
 VERSION I08294.1 GI:588994
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 9185)
 AUTHORS Houghton, M., Choo, Q.-L., and Kuo, G.
 TITLE NANAV diagnostics and vaccines
 JOURNAL Patent, EP 0388232; Al-1-19-SEP-1990;
 FEATURES
 Location/Qualifiers
 1..9185
 /organism="unknown"
 /mol_type="unassigned DNA"


```

Db      5375  GTGCTCTTGTCGGGAAGCGGCAATCATACCTGACAGGAAGTCTCTTACCGAGAGTTC 5434
Qy      681  AspGluMetGluGluCys 686
Db      5435  GATGAGATGGAAGAGTGC 5452

RESULT 11
LOCUS      BD091382                9185 bp    DNA    linear    PAT 27-AUG-2002
DEFINITION HCV cultivation method in eucaryotic cells.
ACCESSION  BD091382
VERSION    BD091382.1 GI:22636993
KEYWORDS  JP 2001314192-A/3.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 9185)
AUTHORS   Weiner,A.J., Steimer,K.S. and Houghton,M.
TITLE     HCV cultivation method in eucaryotic cells
JOURNAL   Patent: JP 2001314192-A 3 13-NOV-2001;
           CHIRON CORP
COMMENT   OS Homo sapiens (human)
           PN JP 2001314192-A/3
           PF 15-MAR-2001 JP 2001075114
           PR 25-AUG-1989 US 398667
           PI AMY J WEINER, KATHELYN S STEIMER, MICHAEL HOUGHTON PC
           C12N15/09, C12N5/10, C12N7/00, C12N1/93, C12N15/00, PC
           C12N5/00
           CC HCV cultivation method in eucaryotic cells
           FH Key
           FT source
           FT Location/Qualifiers
           1..9185
           /organism="Homo sapiens (human)"
           /db_xref="taxon:9606"

FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 3,36e-200 Length: 9185
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservatives: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x BD091382 (1-9185)
Qy      1  MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3395  CTGGCGCCCATCACGGCGTACGCCAGACAGAGGGGCGCTCTAGGGTGCAATACACC 3454
Qy      21  SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      3455  AGCTTAAGTGGCGGGGCAAAACCAACAGTGGAGGTGAGGTCCAGATTGTGTCACTGCT 3514
Qy      41  AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db      3515  GCCCAACCTTCCTGGCAACGTGATCAATGGGGTGTGCTGGAGTGTCTACCAACGGGGCC 3574
Qy      61  GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      3575  GGAACGAGGACCATCGGTACCCCAAGGTCCTGTATCCAGATGATATCAATGATAGAC 3634
Qy      81  GlnAspLeuValGlyTyrProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      3635  CAAGACCTTGTGGGTGGCGCGCTCCGCAAGGTAGCGGCTCAATTGACACCGCTGCATTGC 3694
Qy      101  GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120

```

```

Db      3695  GGCCTCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTGTCATTCCCGTCGGCGGGCGG 3754
Qy      121  GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db      3755  GGTGATAGCAGGGGCGAGCTCTGTGCGCCCGGGCCCATTTCTTACTTGAAGGGCTCTCTCG 3814
Qy      141  GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db      3815  GGGGGTCCGCTGTTGTGCCCCGGGGCACGCCGTGGGCATATTAGGGCCCGCGGTGTGC 3874
Qy      161  ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db      3875  ACCCGTGGAGTGGCTTAAGCGGTGAGCTTTATCCTCTGTGGAGAACCTAGACAAACCATG 3934
Qy      181  ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db      3935  AGGTCCCGGCTGTTACCGGATAACTCTCTCCACCAGTATGTCGCCAGAGTTCACAGTG 3994
Qy      201  AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db      3995  GCTCACTCTCCATGCTCCACAGCGCGGCAAAAGCAACCAAGGTCCCGGCTGCATATGCA 4054
Qy      221  AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db      4055  GCTCAGGCTTATAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGT 4114
Qy      241  AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db      4115  GCTTACATGTCCAAAGGCTCATGGATCGATCTTAACATCAGACCGGGGTGAGAACAAAT 4174
Qy      261  ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db      4175  ACCACTGGCAGCCCCCATCACCTACTCCACCTACCGCAAGTTCCTTGCACGGCGGTGC 4234
Qy      281  SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db      4235  TCGGGGGCGCTTATGACATAATAATTGTGACAGTGCCTCCACGGATGCCACATCC 4294
Qy      301  IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
Db      4295  ATCTTGGGCATCGGCATCTCTTTACCAACGACAGACTGCGGGGGCGAGACTCGTGTGTG 4354
Qy      321  LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      4355  CTCGCCACCGCCACCCCTCCGGGCTCCGTCCTCACTGTGCCCATCCCAACATCGAGAGTT 4414
Qy      341  AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      4415  GCTCTGTCCACGACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 4474
Qy      361  LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db      4475  AAGGGGGGGAGACATCTCATCTTCTGTCATCAAGAGAGAGTGCAGCAACTGCCGCA 4534
Qy      381  LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      4535  AAGCTGGTCCGATTTGGGCATCAATGCCGTGGCTACTACCGCGGTCTTGAAGTCCGTC 4594
Qy      401  IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      4595  ATCCCGACACGGCGGATGTTGTGCTGTGCAACCGATGCCCTCATGACCGGTATACC 4654
Qy      421  GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      4655  GGCAGCTTCGAGTCTGGTGATAGATGCAATACGTGTGTGCACCCACAGACTCGATTTCAGC 4714
Qy      441  LeuAspProThrPheThrIleGluThrIlePheLeuProGlnAspAlaValSerArgThr 460
Db      4715  CTTGACCTTACTCTTACCATTTGAGACAAATCACGTCCTCCCGAGGATGCTGCTCCCGACT 4774
Qy      461  GlnArgAspGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db      4775  CAACCTCGGGCAGGACTGGCAGGGGAAGCCAGGCATCTACAGATTGTGGCACCGGGG 4834

```

Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4835 GAGGCCCTCCGGCATGTTGCGATCGTCTCTGTGAGTGCATGACGAGGCTGT 4894
Qy 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4895 GCTTGGTATGAGCTACGCCCGCGAGACTACAGTTAGGTAGGTACATGACACACC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrPheTyrGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGGCTCCCGGTGTCAGGACCATCTTGAATTTTGGGAGGCGCTCTTTACAGGCGCTC 5014
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5074
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 5075 CTGGTAGCGTACCAAGCCCGTGTGGCTAGGCTCAAGCCCTCCCCCATCGTGGGAC 5134
Qy 581 GlnMetTrpLysCysIleuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGAAGTGTGTGATTCGCTCAAGCCACCCCTCATGGGCCAACCCCTGCTA 5194
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGCGCTGTTCAGATGAATCAACCTGACGCCACCCAGTCAACCAATACATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrProValLeuValGlyGly 640
Db 5255 ATGACATGTCATGTGGCGGACCTGGAGGTGCTCACGAGCACCTGGGTGCTCGTGGCGG 5314
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGTGCTGTGGCGGATGTGCTGTCAACAGCTGCGGTGTCATAGTGGCAGG 5374
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTGCTGTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCAGAGTTC 5434
Qy 681 AspGluMetGluGluCys 686
Db 5435 GATGAGATGAAGAGTGC 5452

RESULT 12
ARI166930
LOCUS ARI166930 9379 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6284249.
ACCESSION ARI166930
VERSION ARI166930.1 GI:16243325
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 9379)
AUTHORS Barban,V.
TITLE Fusion polypeptide having the C protein and E1 protein of hepatitis C virus
JOURNAL C virus
Patent: US 6284249-A 1 04-SEP-2001;
FEATURES Location/Qualifiers
source 1..9379
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3.45e-200 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x ARI166930 (1-9379)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3395 CTGGCGCCCATCACGGCGTAGCCCCAGCAGCAAGGGGCGCTCTAGGGTGCAATATCACC 3454
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3455 AGCCTAACTGGCGCGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 3514
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db 3515 GCCCAACCTTCTGGCAACGTCATCAATGGGTGTCTGGACTGTCTACACGGGCGC 3574
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3575 GGAACGAGGACCATCGCTCACCCAAAGGTCTCTGTCACTCAGATGTATACCAATGTAGAC 3634
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3635 CAAGCACTTGTGGGTGGCGCTCCGCAAGGTAGCGCTCATTGACACCTGCACTTGC 3694
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 3695 GGCTCTCGGACCTTTACCTGTGTACAGAGGACGCGCATGTCACTCCCGTGGCGCGGG 3754
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3755 GGTGATAGCAGGGGAGGCTGTGTGCGCCGCGCCCATTTCTACTTGAAGGCTTCTCG 3814
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3815 GGGGTGCGTGTGTGCGCGGCGCACGCGTGGGCATATTTAGGCGCGGTGTGC 3874
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3875 ACCCTGGAGTGGCTAAGCGGTGACATTATCCCTGFGAGAACTAGAGCAACACCATG 3934
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 3935 AGGTCCCGGTGTTACCGGATACTCTCTCCACAGTAGTGCCTCCAGAGTTCAGGTG 3994
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3995 GGTCACTCCATGTCTCCACAGGCGGCAAGCAACCAAGGTCCCGGCTGCATATGCA 4054
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4055 GCTCAGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGCACTGGGCTTGGT 4114
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCCAGGCTCATGGGATCGATCCTAAATCAGGACCGGGGTGAGAACAAAT 4174
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
Db 4175 ACCACTGGAGCGCCCATCACGTACTCCACCTACGCAAGTTCCTTGGCAGCGCGGTGC 4234
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGCGCTTATGACATAATAATTGTGACGAGTCCACTCCACGGATGCCACATCC 4294
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnIleGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTTGGGCGATCGGCACTGTCTTGTACCAAGCAGAGACTGCGGGGCGAGACTGGTGTG 4354
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTCGCCACCGCCACCCCTCCGGGTCCGTCACTGTGCCCCCATCCCAACATCGAGAGGT 4414
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360

Db	4415	GCTCTGTCACCA	CGGAGAGATCCCTTTT	TACGGCAGGCTATCC	CCCTCGAAGTAATC	4474
Qy	361	LysGlyGlyArgHisLeu	IlePheCysHisSerLys	LysLysCysAspGluLeu	AlaAla	380
Db	4475	AAGGGGGGAGACAT	CTCATCTTCTGTCTCAT	TCAAGAGAGAGTGG	ACGACGACCTCGCCGCA	4534
Qy	381	LysLeuValAlaLeu	GlyValAsnAlaValala	TyrTyrArgGlyLeu	AspValSerVal	400
Db	4535	AAGCTGCTCGCAT	TGGGCATCAATGCGT	GGCCTACTACCGCGGT	CTTGACGTGTCCGCTC	4594
Qy	401	IleProThrSerGly	AspValValAlaAla	ThrAspAlaLeuMet	ThrGlyPheThr	420
Db	4595	ATCCGACACAGCG	CGATGTTTCGTCTG	TGGCAACCGATGCC	CTCATGACCGGCTATACC	4654
Qy	421	GlyAspPheAspSer	ValIleAspCysAsnThr	CysValThrGlnThrVal	AlaAspPheSer	440
Db	4655	GGCGACTTCGACT	TCGGTGATAGACTG	CAATACGTGTGTCA	CCACAGACAGTCGATTTCAGC	4714
Qy	441	LeuAspProThrPhe	ThrIleGluThrIle	ThrLeuProGlnAsp	AlaValSerAsnThr	460
Db	4715	CTTGACCCTACCT	TTACCATTTGACACAA	TACCGTCCCGCAGG	ATGCTGTCTCCCGCACT	4774
Qy	461	GlnArgArgGlyArg	ThrGlyArgGlyLys	ProGlyIleTyrArgPhe	ValAlaProGly	480
Db	4775	CAACGTCGGGGC	AGACTGCGAGGGGA	ACGAGCATCTACAG	ATTGTGGCACCGCGG	4834
Qy	481	GluArgProSerGly	MetPheAspSerSer	ValLeuCysGluCys	TyrAspAlaGlyCys	500
Db	4835	GAGCGCCCTCCG	GCATGTTTCGACTCG	CTCGTCTGTGAGTG	CTATGACGAGGCTGT	4894
Qy	501	AlaTrpTyrGluLeu	ThrProAlaGluThr	ThrValArgLeuArg	AlaTyrMetAsnThr	520
Db	4895	GCTTGTGATGAG	CTCACGCCCGCGAG	CTACAGTTAGGCTAC	GAGCGGTACATGAA	4954
Qy	521	ProGlyLeuProVal	CysGlnAspHisLeu	GluPheTrpGluGly	ValPheThrGlyLeu	540
Db	4955	CCGGGGCTTCCG	GTGCCAGGACATCT	TGAAATTTTGGAG	GGCGCTCTTTACAGG	5014
Qy	541	ThrHisIleAspAla	HisPheLeuSerGln	ThrLysGlnSerGly	GluAsnLeuProTyr	560
Db	5015	ACTCATATAGAT	GCCCATTTCTATCC	CAGAAAGCAGAG	TGGGAGAACCTTCCTT	5074
Qy	561	LeuValAlaTyrGln	AlaThrValCysAla	ArgAlaGlnAlaPro	ProProSerTrpAsp	580
Db	5075	CTGCTAGCGTACC	AAGCACCGTGTGCG	TAGGGCTCAAGC	CCCTCCCATCGTGGAC	5134
Qy	581	GlnMetTrpLysCys	LeuIleArgLeuLys	ProThrLeuHisGly	ProThrProLeuLeu	600
Db	5135	CAGATGTGAAGT	GTGTTGATTCGGCT	CAAGCCACCTCTCA	TGGGCCAAACACCCCT	5194
Qy	601	TyrArgLeuGlyAla	ValGlnAsnGluVal	ThrLeuThrHisPro	ValThrLysTyrIle	620
Db	5195	TACAGACTGGCG	CGCTGTTCAGATGA	ATAATCACCTGAC	CGCACCCAGTCAACAA	5254
Qy	621	MetThrCysMetSer	AlaAspLeuGluVal	ValThrSerThrTrp	ValLeuValGlyGly	640
Db	5255	ATGACATGCTATG	TCGGCGGCACCTG	AGGTGGTTCACGAG	CACTGGGTGCTCGT	5314
Qy	641	ValLeuAlaAlaLeu	AlaAlaTyrCysLeu	SerThrGlyCysVal	ValIleValGlyArg	660
Db	5315	GTCTGTGCTGCT	TGGCGCGGTATTG	CTGTCAACAGCGT	TCGTGGTTCATAGTGG	5374
Qy	661	IleValLeuSerGly	LysProAlaIleIle	ProAspArgGluVal	LeuTyrArgGluPhe	680
Db	5375	GTCGTCTGTGTC	CGGGGAACCGGCA	TACATACCTGAC	AGGAAGTCTCTTACC	5434
Qy	681	AspGluMetGlu	GluCys	686		
Db	5435	GATGAGATGGA	AGAGTGC	5452		
RESULT 13						
AR301300						
LOCUS	AR301300	9379 bp	DNA	linear	PAT 12-JUN-2003	

DEFINITION	Sequence 1 from patent US 638123.
ACCESSION	AR301300
VERSION	AR301300.1 GI:31689076
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 9379)
TITLE	Barban, V.
JOURNAL	Vaccine composition for preventing or treating hepatitis C
FEATURES	Patent: US 638123-A 1 25-MAR-2003;
source	Location/Qualifiers
	1..9379
	/organism="unknown"
ORIGIN	/mol_type="genomic DNA"
Alignment Scores:	
Pred. No.:	3,456-200 Length: 9379
Score:	3574.00 Matches: 672
Percent Similarity:	99.42% Conservative: 10
Best Local Similarity:	97.96% Mismatches: 4
Query Match:	98.78% Indels: 0
DB:	Gaps: 0
US-09-930-591-2 (11-686) x AR301300 (11-9379)	
Qy	1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db	::: 3395 CTGGCGCCCATCAGCGGTACGCCCGACACAAAGGGGCGCTCTCTAGGGTGCAATAACACC 3454
Qy	21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db	3455 AGCTAACTCGCGGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTCAACTGCT 3514
Qy	41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db	3515 GCCAAACCTTCTCGCAAGTCATCAATGGGGGTGTGCTGAGCTGCTACCAAGGGGCC 3574
Qy	61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db	3575 GGAACGAGGACCATCGGTCACCAAGGTCCTGTCTATCCAGATGTATACCATGTAGAC 3634
Qy	81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db	3635 CAAGACCTTGTGGCTGGCGCGCTCCGCAAGGTAGCGGCTCAITGACACCCCTGCCTTGC 3694
Qy	101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db	3695 GGTCTCTCGACCTTTACCTGTGTCCCGCGGGGACGCGCGTGGCATATTCCGTCGCCGGCGG 3754
Qy	121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db	3755 GGTGATAGAGGGGACGCTGTGTGCCCGGCCCATTTCTACTTGAAGGCTCTCTCG 3814
Qy	141 GlyClyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db	3815 GGGGGTCCGTGTGTGTCGCCCGGGGACGCGCGTGGGCATATTAGGGCGCGGGTGTGC 3874
Qy	161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db	3875 ACCCGTGGATGGCTAAGCGCGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACCATG 3934
Qy	181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db	3935 AGGTCCCGGTGTTCCGGATAAATCTCTCTCCACCATGTATGTGCCCCAGAGCTTCCAGGTG 3994
Qy	201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db	3995 GCTCACCTCATGCTCCACAGCAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA 4054
Qy	221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIleThrMetGlyPheGly 240

Db 4055 GCTCAGGGCTATAAGCTGTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCGAAGCTCATGGATCGGATCTTAACATCAGGACCGGGGTGAGACAAT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACTGGCAGCCCATCAGTACTCCACTAGCGCAAGTTCCTTGCCGACGGCGGTG 4234
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGGGCGCTTATGACATAATTTGTGACGAGTGCATCCACGGATGCCATCC 4294
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTTGGGCAATCGGCACTGCTTGTACCAAGCAGAGACTCGGGGGCGAGACTGGTGTG 4354
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTGCCACCGCCACCCCTCGGGCTCGCTCACTGTGCCCATCCCAACATCGAGGAGGT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTCTCCACCGGAGAGATCCCTTTTACGCAAGGCTATCCCCCTCGAAGTAATC 4474
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaIa 380
Db 4475 AAGGGGGGAGACATCTCTCTCTGTCTTCAAGAAGAGTCCGACAACTCGCGCA 4534
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4535 AAGCTGTGCGATGGGCACTCAATCGGTGGCTACTACCGGGTCTTACGGTGTCCGTC 4594
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCGACGAGCGGATGTGTGTGTGTGGCAACCGATGCCCTCATGCCGTATACC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GCGGACTTCGACTCGGTGTAGACTCAATACGTGTGTCCACGACAGCTCGATTTCAG 4714
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCTTACCTTCCATTGAGACATACGCTCCCCAGATGCTGTCTCCCGCAT 4774
QY 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4775 CAACGTGCGGGCAGGACTGCGAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 4834
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4835 GAGCGCCCTCCGGCATGTTCGACTCGTCCGTCTCTGTGAGTGTCTATGACGAGGCTGT 4894
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4895 GCTTGGTATGAGTCAAGCCCGCGAGACTACAGTTAGGCTACAGCGGTACATGACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGGCTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGCTTTACAGGCCCTC 5014
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580
Db 5075 CTGCTAGCGTACCAAGCCCGTGTGGCTAGGGCTCAAGCCCTCCCCCATCGTGGGAC 5134
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCCAACCCCTCCATGGGCCAACACCCCTGCTA 5194

QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGCGCTGTTCAGATGAATCACCTTCGACCCAGTCCACCAATATCATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5255 ATGACATGATGTCTGGCGACCTGGAGTGTGTACAGAGCACCTGGTGTCTGTTGGGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTTGGCTGTCTTGGCGCGTATTGCTGTCAACAGCTCGGTGTCATAGTGGGAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTCTGTCTGTCTGGGAAGCGGCAATCATACCTGACAGGAAGTCTCTACCGAGATTC 5434
QY 681 AspGluMetGluGluCys 686
Db 5435 CATGAGATGGAAGAGTGC 5452

RESULT 14
AR176483 9401 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 9 from patent US 6312889.
DEFINITION AR176483
ACCESSION AR176483
VERSION AR176483.1 GI:17918838
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 9401)
AUTHORS Houghton,M., Choo,Q.-J. and Kuo,G.
TITLE Combinations of hepatitis c virus (HCV) antigens for use in
immunassays for anti-HCV antibodies
JOURNAL Patent: US 6312889-A 9 06-NOV-2001;
FEATURES Location/Qualifiers
source 1..9401
/organism="unknown"
/mol_type="unassigned DNA"

Alignment Scores:
Pred. No.: 3,46e-200 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AR176483 (1-9401)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGCGCCCATCACGGGTACGCCACAGACAGAGGGGCTCTTAGGGTGCATATATCACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCTTAATGCGCGGACAAACCAAGTGGAGGTTGAGTCCAGATGTGTCACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAAACTTCTCTGGCAACGTGCATCAATGGGTGTGTCTGACTGTCTACACGGGCCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAACGAGACATCGGCTCACCAAGGGTCTCTGTCTATCCAGATGTATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGGTGGCGCTCCGCAAGGTAGCGCTCATTGACACCCCTGACATGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 15, 2004, 15:46:56 ; Search time 133 Seconds
(without alignments)
2862.378 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPITAYAAQTGCLGCIIT.....PALIPDREVLVREFDEMEEC 686

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US0930591/runat 15092004.164647 7129/app query.fasta 1.839
-DB=Issued Patents NA -QPM=fastap -Suffix=nni -NIMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0930591 @CEN 1.1 59 @runat 15092004.164647 7129 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DEEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCBUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3574	98.8	7310	3	US-08-444-818-74
2	3574	98.8	9379	3	Sequence 74, Appl
3	3574	98.8	9379	4	Sequence 1, Appl
4	3574	98.8	9401	1	US-09-916-359-1
5	3574	98.8	9401	1	Sequence 9, Appl
6	3574	98.8	9401	1	US-07-910-760-9
7	3574	98.8	9401	4	US-08-440-519-9
8	3574	98.8	9401	4	Sequence 9, Appl
9	3574	98.8	9401	4	US-08-440-549-9
10	3574	98.8	9401	4	Sequence 25, Appl
11	3574	98.8	9401	4	US-08-823-895A-25
12	3574	98.8	9401	4	Sequence 55, Appl
13	3574	98.8	9401	4	US-08-444-818-65
14	3574	98.8	9401	4	Sequence 122, Appl
15	3574	98.8	9401	4	US-08-444-818-88
16	3574	98.8	9401	4	Sequence 88, Appl
17	3574	98.8	9401	4	US-08-444-818-122
18	3574	98.8	9401	4	Sequence 123, Appl
19	3574	98.8	9401	4	US-08-444-818-123
20	3574	98.8	9401	4	Sequence 2, Appl
21	3574	98.8	9401	4	US-09-881-239-2

13	3565	98.5	2058	4	US-09-881-654-1	Sequence 1, Appl
14	3565	98.5	9887	3	US-08-444-818-137	Sequence 137, App
15	3555	99.3	5360	3	US-08-444-818-53	Sequence 53, Appl
16	3551	98.1	9646	3	US-08-811-566-1	Sequence 1, Appl
17	3551	98.1	9646	4	US-09-034-756-1	Sequence 5, Appl
18	3551	98.1	12980	3	US-08-811-566-5	Sequence 5, Appl
19	3551	98.1	12980	4	US-09-034-756-5	Sequence 2, Appl
20	3544	98.0	9599	3	US-09-014-416-2	Sequence 6, Appl
21	3544	98.0	9599	3	US-09-014-416-6	Sequence 1, Appl
22	3538	97.8	9401	2	US-08-432-693-1	Sequence 19, Appl
23	3538	97.8	9416	4	US-08-811-566-19	Sequence 19, Appl
24	3538	97.8	9416	4	US-09-034-756-19	Sequence 176, App
25	3535	97.7	9379	3	US-08-444-818-176	Sequence 9, Appl
26	3532	97.6	9401	5	PCT-US91-02225-9	Sequence 26, Appl
27	3504	96.8	9416	4	US-08-823-895A-26	Sequence 13, Appl
28	3504	96.8	9416	4	US-10-104-966-13	Sequence 10, Appl
29	3429	94.8	7989	4	US-09-539-601-10	Sequence 7, Appl
30	3429	94.8	8001	4	US-09-539-601-7	Sequence 4, Appl
31	3426	94.7	8637	4	US-09-539-601-4	Sequence 13, Appl
32	3426	94.7	8649	4	US-09-539-601-13	Sequence 1, Appl
33	3426	94.7	11076	4	US-09-539-601-1	Sequence 22, Appl
34	3423	94.6	8001	4	US-09-539-601-22	Sequence 11, Appl
35	3420	94.5	6039	1	US-08-324-977-11	Sequence 11, Appl
36	3420	94.5	6039	2	US-08-384-616-11	Sequence 11, Appl
37	3420	94.5	6039	2	US-08-904-686A-11	Sequence 11, Appl
38	3420	94.5	6039	3	US-09-315-850-11	Sequence 35, Appl
39	3420	94.5	7863	1	US-08-324-977-35	Sequence 35, Appl
40	3420	94.5	7863	2	US-08-384-616-35	Sequence 35, Appl
41	3420	94.5	7863	2	US-08-904-686A-35	Sequence 35, Appl
42	3420	94.5	7863	3	US-08-315-850-35	Sequence 35, Appl
43	3420	94.5	7917	1	US-08-324-977-31	Sequence 31, Appl
44	3420	94.5	7917	2	US-08-384-616-31	Sequence 31, Appl
45	3420	94.5	7917	2	US-08-904-686A-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-08-444-818-74
; Sequence 74, Application US/08444818
; Patent No. G750087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; Sequence 74, Floppy disk
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:
 LENGTH: 7310 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3..7310
 US-08-444-818-74

Alignment Scores:

Pred. No.: 0 Length: 7310
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-74 (1-7310)

QY 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1728 CTGGCGCCCATCAGGGGTACGCCAGCAGACAGAGGGGCTCTTAGGGTGCATAATCACC 1787
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 1788 AGCCTAACTGGCCGGGCAAAACCAAGTGAGGGGTGAGGTCCAGATTGTCAACTGCT 1847
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 DB 1848 GCCCAAACTTCCTGGCAACGTGATCAATGGGGGTGCTGGACTGCTACCAACGGGGCC 1907
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 1908 GGAACGAGGACCATCGGTCAACCAAGGTCCTGTATCCAGATGTATACCAATGTAGAC 1967
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 1968 CAAGACCTTGTGGCTGGCCCGCTCCCAAGGTAGCGGCTCATGTACACCCCTGCACCTTC 2027
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 2028 GCCTCTCGACCTTTACCTGGTCAAGGACGCGCATGCTATCCCGTGGCCCGGGCGG 2087
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 2088 GGTGATAGCAGGGGCGAGCTGCTGCGCCCGGGCCCATTTCTACTTTGAAAGGCTCCTCG 2147
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 2148 GGGGGTCCGCTGTGTGCCCCCGGGGACGCGGTGGGCAATTTAGGGCCGCGGTGTGC 2207
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 2208 ACCCGTGGAGTGGCTAAGGGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACCATG 2267
 QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
 DB 2268 AGGTCCCGGGTTCACGGTAATCTCTCCACAGTAGTGGCCAGAGCTTCCAGGGT 2327
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 2328 GCTCACCTCCATGTCTCCACAGGCGAGCGGCAAAAGACCAAGTCCCGGCTGCATATGCA 2387
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIleThrMetGlyPheGly 240
 DB 2388 GCTCAGGGCTAAGGGTGTAGTACTCAACCCCTCTGTCTGCACACACTGGGCTTTGGT 2447
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 2448 GCTTACATGTCGAAGGCTCATGGGATCGATCTAATCAATCAGGACCGGGGTGAGACAATT 2507

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 2508 ACCACTGGCAGCCCATCAGTACTCCACCTACGCAAGTTCTTTCGCGAGCGGGGTGC 2567
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 2568 TCGGGGGGCGCTTATGACATAATATTTGTGACGAGTGCCTCCACCGATGCCACATCC 2627
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 2628 ATCTTGGGCATCGGCATCTCTTACCAGCAGAGACTGGGGGGGAGACTGGTTGTG 2687
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluVal 340
 DB 2688 CTCGCCACCGCCACCCCTCCGGGCTCCGCTCACTGTGCCCATCCCAACATCGAGAGGTT 2747
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 DB 2748 GCTCTGTCCACCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAAATC 2807
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 DB 2808 AAGGGGGGAGACATCTCATCTCTGTCTATTCAAGAGAGAGTGCAGCAACTCGCGCA 2867
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 2868 AAGCTGGTTCGATTTGGGCATCAATGCCGTGCTACTACCGCGGTCTTGACGTGTCGTC 2927
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 2928 ATCCCGACACCGGGGATGTTGCTGCTGGCAACCGATGCCCTCATGACCGGTATACC 2987
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 2988 GCGCATCTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTTCGATTCAGC 3047
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 3048 CTTGACCTTACTTCCACATTGAGACATCAGCTCCCGAGGATGCTGTCTCCCGCACT 3107
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 DB 3108 CAACGTCCGGGCGAGACTGGCAGGGGGAAGCAGGCATCTACAGATTGTGGCACCGGG 3167
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 DB 3168 GAGCGCCCTCCGGCATGTTGACTCGCTCCCTCTGTGAGTGTATGACGCGAGGCTGT 3227
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 3228 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTAGGCTACGAGCGGTACATGAACACC 3287
 QY 521 ProGlyLeuProValCysGlnAspHisLeuIlePheTrpGluGlyValPheThrGlyLeu 540
 DB 3288 CGGGGGCTTCCCGTGTGCGCAGGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGCTC 3347
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 DB 3348 ACTCATATAGATGCCACTTCTATCCAGACAAGCAGAGTGGGGAGAACCTTCTTCTTAC 3407
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerThrAsp 580
 DB 3408 CTGTGTAGCGTACCAAGCCACCGCTGTGGCTAGGCTCAAGCCCTCCCGCATCTGGGAC 3467
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 3468 CAGATGTGGAAGTGTGTGATTCGCTCCAGCCACCTCCATGGGCCCAACACCTCTGCTA 3527
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 DB 3528 TACAGACTGGGGCTGTTCAGAAATCAATCACCCTGACGACCCAGTCCACCAATACATC 3587
 QY 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640

```

Db      3588  ATGACATGCATGTGGCGGCACCTGGAGTGGTCTACGAGCAGCCTGGGGTGGCTGGTGGCGGC 3647
Qy      641  ValLeuAlaAlaLeuAlaAlaTyrcysLeuSerThrGlyCysValValIleValGlyVarg 660
Db      3648  GTCCTGCTGTCTTGGCGGGTATTGGCTGTCAACAGGCTCGGTGGTCATAGTGGGAGG 3707
Qy      661  IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrcysGluPhe 680
Db      3708  GTCGTCTTGTCCGGGAAGCCGGCAATCATACCTGACAGGAAGTCCCTTACCGAGAGTTC 3767
Qy      681  AspGluMetGluGluCys 686
Db      3768  GATGAGATGGAAGAGTGC 3785

RESULT 2
US-09-388-874-1
; Sequence 1, Application US/09388874
; Patent No. 6284249
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PNC97-03A
; CURRENT APPLICATION NUMBER: US/09/388,874
; CURRENT FILING DATE: 1999-09-02
; EARLIER APPLICATION NUMBER: PCT/FR98/00448
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 97/02,887
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320)...(9352)
US-09-388-874-1

Alignment Scores:
Pred. No.: 0 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-09-388-874-1 (1-9379)
Qy      1  MetaAlaProIleThrAlaTyrcysAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3395  CTGGCGCCCATCACGGCGTAGCCCGCAGCAGACAGAGGGCCCTCTTAGGGTGCATAATCACC 3454
Qy      21  SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      3455  AGCCTAACTGGCCGGGCAAAAACCAATGGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3514
Qy      41  AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysrtpThrValTyrcysGlyVala 60
Db      3515  GCCCAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGAGTGTCTACCACGGGCCC 3574
Qy      61  GlyThrArgThrIleAlaSerProGlyGlyProValIleGlnMetTyrcysValAsp 80
Db      3575  GGAACGAGGACCATCGCGTACCACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 3634
Qy      81  GlnAspIleValGlyTrpProAlaProGlnGlyAlaArgSerIleuThrProCysThrCys 100
Db      3635  CAAGACTTGTGGGTGGCGCGCTCCGCAAGTAGCCGCTCAATTGACACCCCTGCACATTGC 3694
Qy      101  GlySerSerAspLeuTyrcysValIleThrArgHisAlaAspValIlePrcValArgArg 120

```

QY 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 4835 GAGCGCCCTCCGCGCATGTCGACTCGTCGCTCTGTGAGTGCTATGACGAGGCTGT 4894
 QY 501 AlaTrpTyrGluLeuThrProAlaGluThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 4895 GCTTGGTATGAGCTACGCGCCGCGAGACTACAGTTAGGTAGCGTACATCAACACC 4954
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 4955 CCGGGCTTCCGCTGTCAGACCACTTGAATTTTGGGAGGCGCTTTACAGGCTC 5014
 QY 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 5015 ACTCATATAGTGCCTTCTATCCAGCAAGAGAGAGTGGGAGAACCTTCTCTTAC 5074
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 5075 CTGTAGCGTACCAAGCCAGCTGTGGCTAGGCTCAAGCCCTCCCATCGTGGGAC 5134
 QY 581 GlnMetTrpLysCysLeuLeuArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 5135 CAGATGTGAAGTGTGTTGANTCGCTCAAGCCCAACCTCCATGGCCCAACACCCCTGCTA 5194
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 5195 TACAGACTGGGCGCTGTTCAGATGAATACCTGACGACCCAGTCACCAATACATC 5254
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 5255 ATGACATGTCATGTCGGCGACCTGAGGTGTTCACAGACACCTGGGTGCTCGTTGGCGC 5314
 QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 5315 GTCTGGCTGCTTTGGCGCGTATGCTGTCAACAGGCTCGGTGTCATAGTGGCGAG 5374
 QY 661 IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuTyrArgGluPhe 680
 Db 5375 GTCTGCTTGTCCGGAGCCGCGCAATCATACCTGACAGGAGTCTCTACCGAGAGTTC 5434
 QY 681 AspGluMetGluGluCys 686
 Db 5435 GATGAGATGGAAGATGC 5452

RESULT 3
 ; Sequence 1, Application US/09916359
 ; Patent No. 6538123
 ; GENERAL INFORMATION:
 ; APPLICANT: Veronique Barban
 ; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
 ; TITLE OF INVENTION: TREATING C HEPATITIS
 ; FILE REFERENCE: PWC97-03A
 ; CURRENT APPLICATION NUMBER: US/09/916,359
 ; CURRENT FILING DATE: 2001-07-26
 ; PRIOR APPLICATION NUMBER: 09/388,874
 ; PRIOR FILING DATE: 1999-09-02
 ; PRIOR APPLICATION NUMBER: 97/02,887
 ; PRIOR FILING DATE: 1997-03-06
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 9379
 ; TYPE: DNA
 ; ORGANISM: Virus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (320)...(9352)
 ; US-09-916-359-1

Alignment Scores:
 Pred. No.: 0 Length: 9379

Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 4 Gaps: 0
 US-09-930-591-2 (1-686) x US-09-916-359-1 (1-9379)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 3395 CTGGCGCCCATCACGCGCTACGCCAGCAGACAAGGGCCTCTAGGTGTCATAATCAC 3454
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 3455 AGCCTAACTGGCCGGGACAAAACCAAGTGAGGTGAGGTCCAGATTTGTCAACTGCT 3514
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 3515 GCCCAACCTTCTGGCAACGTCATCAATGGGTGTGCTGACTGTCTACACGGGGCC 3574
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 3575 GGAACGAGGACCATCGCGTCAACCAAGGTCTCTGTATCCAGATGTATACCAATGTAGAC 3634
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 3635 CHAGACCTTGTGGGTGGCCCGCTCGCAAGGTAGCGCTCATTTGACACCTCGTCTGCT 3694
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 Db 3695 GGCCTCTCGGACCTTTTACCTGTCACGAGGACCCGATGTCAATTCGCTGCGCGCGCG 3754
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 3755 GGTGATAGCAGGGGAGCTGCTGTGCGCCCGCGCCATTTCTACTTGAAGGCTCTCTCG 3814
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 3815 GGGGTGCTGCTGTGTGTCGCCCGCGGCGCACGCCGTGGGCATATTTAGGGCGCGGTGTC 3874
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
 Db 3875 ACCCGTGGAGTGGCTAAGCGGTGACTTTATCCTGTGGAGAACCTTAGAGACCAACATG 3934
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 3935 AGGTCCCGGTGTTTACCGGATACTCTCTCCACAGTAGTGTGCCCCAGAGCTTCCAGGTG 3994
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 3995 GGTCACTCTCCATGCTCCACAGGCGCGCAAGAGCCCAAGAGTCCCGGTGTCATATGCA 4054
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 4055 GCTCAGGCTTATAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGGT 4114
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4115 GCTTACATGTCCAAGGCTCATGGGATCATCTTAACATCAGAGCCGGGGTGGAGAACAT 4174
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 4175 ACCACTGGACGCCCATCAGTACTTCCACCTACGGCAAGTCTCTTGGCGAGGGGGTGC 4234
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4235 TCGGGGGCGCTTATGACATAATAATTGTGACAGTGCCACTCCACGGATGCCACATCC 4294
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 4295 ATCTTTGGGCATCGGCACCTGCTCTTGACCAAGAGACTGGGGGGGCGAGACTGGTGTG 4354
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340

Db 5435 GATGAGTGAAGAGTC 5452

RESULT 4
US-07-910-760-9
Sequence 9, Application US/07910760
Patent No. 5683864
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07910,760
FILING DATE: 07-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9401 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 342..9374
FEATURE:
NAME/KEY: misc_feature
LOCATION: 366
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Arg."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc feature

Db 5435 GATGAGTGAAGAGTC 5452

4355 CTCGCCACCGCCACCCTCCGGGTCTCCCTACTGTGCCCACTCCCAACATCGAGGAGTT 4414

341 AlaLeuSerThrThrGlyGluIleProPheTyrglyLysAlaIleProLeuGluAlaIle 360

4415 GTCTGTCTCACCGGAGATCCCTTTTACGGCANGGTATCCCCCTCGAAGTAATC 4474

361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380

4475 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAAGAAGAAGTGCGACAACTCGCGCA 4534

381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrtyrArgGlyLeuAspValserVal 400

4535 AGCTGTCGCATTTGGGCAATCCCGTGTCTACTACCGCGTCTTTCAGCTTCGTC 4594

401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420

4595 ATCCGACACGCGGGATGTTGTCGTGTCGTCGCAACCGATGCCCTCATGACCGGTATACC 4654

421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

4655 GGCGACTTCGATCGGTGATGACATGCAATAGTGTGTACCCAGACAGTGTATTCAGC 4714

441 LeuAspProThrPheThrileGluThrileThrLeuProGinAspAlaValserArgThr 460

4715 CTTGACCTTACCTTACCATTGAGACAATCAGCTCCCCAGGATGCTGTCTCCGCACT 4774

461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480

4775 CAACGTCCGGGAGGACTGGCAGGGGGAAGCCAGGCACTACAGATTGTGGCACCGGGG 4834

481 GluArgProSerGlyWetPheAspSerValLeuGlyCysGluCysTyrAspAlaGlyCys 500

4835 GAGCGCCCTCCGGCATGTCACCTGCTCCGCTCTGTGAGTGTATGACGCGAGGTGT 4894

501 AlaTrpTyrGluLeuThrProAlaGluThrValArgLeuArgAlaTyMetAsnThr 520

4895 GCTTGTATGAGCTACGCGCCCGGAGACTACAGTAGGCTACGAGCGTACATGAACACC 4954

521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540

4955 CGGGGCTTCCCGTGTGCGAGGACCATCTGAAATTTGGGAGGGCGTCTTACAGGCTC 5014

541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560

5015 ACTCATATAGTCCACCTTCTATCCCAGACAAAGAGAGTGGGGAGAACCTTCTTAC 5074

561 LeuValAlaTyrcGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpasp 580

5075 CTGGTACGCTACCAAGCCACCGTGTGCGTAGGGCTCAAGCCCTCCCGCCATCGTGGAC 5134

581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600

5135 CAGATGTGNAGTGTGTTGATTCGCTCAGCCACCCCTCCATGGGCCACACCCCTGCTA 5194

601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrlle 620

5195 TACAGACTGGGCGCTGTCAGAAATGAATCATCCCTGACGCCACCCAGTACCACAAATACATC 5254

621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640

5255 ATGACATGCAATGTCGGCCACCTGGAGGTGTCACAGACCTGGGTGCTGTGGCGGC 5314

641 ValLeuAlaAlaLeuAlaAlaTyrcLysLeuSerThrGlyCysValValIleValGlyArg 660

5315 GTCCGTGCTGTTGGCGGCTATTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 5374

661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680

5375 GTGCTGCTTGTCCGGGAGCGGCAATCATACCTGACAGGAGTCTCTTACCGAGAGTTC 5434

681 AspGluMetGluGluCys 686

```

/ LOCATION: 2883
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Asn."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3681
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3690
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4167
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Leu."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4323
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4701
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Tyr."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4752
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."

```

```

/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
US-07-910-760-9

Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0

US-09-930-591-2 (1-686) x US-07-910-760-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGCGCCCATCACGGCTAGCCGACAGACAAAGGGCCCTCTAGGGTGTCATAATCAC 3476
QY 21 SerLeuThrGlyAcgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCCTAACTGGCGGACAAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAACCTTCTGGCAACGTGCATCAATGGGTGTGTGGACTGTCTACACGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGACAGGACCATCGCGTCACCAAGGGTCTGTCTCAGATGTATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTAGCGCTCATTCACACCTGACATTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3717 GGCTCTCGGACCTTTACCTGGTCACGAGGACCGCATGTCTCCCGTCCGCGCGCGG 3776
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3777 GGTGATAGCAGGGGAGCGCTGTCTGCCCGCGGCCCATTTCTACTTGAAGGCTCTCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGGTCCGCTGTTGTGCCCGCGGGGACGCCGTGGGCATATTTAGGGCGCGCGTGTGC 3896
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3897 ACCGTGGAGTGGCTAAGCGGTGACTTTATCTCTGTGGAGAACCTTAGACACACCATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGGTCCCGGTGTTTACAGGATAACTCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlySerThrLysValProAlaAlaTyrAla 220
Db 4017 GCTCACCTCCATGTCTCCACAGGACGGGCAAGAACCAAGGTCCTCCGCTGCTATATGCA 4076
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

```

Db 4077 GCTCAGGGCTAAGTGCTAGTACTCAACCCCTCTGTGCTGCACACTGGGCTTTGGT 4136
 Qy 241 AlaTyrMetSerLysAlaHisGlyLeuAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4137 GCTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATT 4196
 Qy 261 ThrThrGlySerProIleThrThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 Db 4197 ACCACTGGGAGGAGCCCATCAGTACTCCACTACGGCAAGTTCCTTGCCACCGCGGGTGC 4256
 Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4257 TCGGGGGGCGCTTATGACATAATAATTGTGACGAGTGCCACTCCACGAGTCCACATCC 4316
 Qy 301 IleGluGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgThrVal 320
 Db 4317 ATCTTGCGCATCGGCATGCTGCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTG 4376
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 4377 CTGCGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCCCATCCCAACATCGAGGAGGT 4436
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle 360
 Db 4437 GCTCTGTCCACCGCAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAAGTAATC 4496
 Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 Db 4497 AAGGGGGGAGACATCTCATCTCTGTCAATCAAGAGAGTGCAGCACTCGCGCA 4556
 Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 4557 AAGCTGTGCTGCTTGGCATTAATGCGTGGCTACTACCGCGCTGTGACGTGTCGCTC 4616
 Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 4617 ATCCCGACCGCGCGATGTGCTGCTGCGCAACCGATGCCCTCATGACCGGCTATACC 4676
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 4677 GCGGACTTCGACTCGGTAGTACTGCAATACGTGTGTACCCAGACAGTCGATTTTCCAGC 4736
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4737 CTTGACCCCTACCTTCACTTACGATGACAAATCAGCTCCCGGATGCTGTCTCCCGCAT 4796
 Qy 461 GluArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 4797 CAAGCTCGGGGAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGG 4856
 Qy 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 4857 GAGCGCCCTCCGGCATGTTTCGACTCGCTCGTCTGTGAGTGCTATGACGCGAGGCTGT 4916
 Qy 501 AlaTyrThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 4917 GCTTGGTATGAGCTACGCGCGCGGAGACATACAGTTAGGCTACGAGGTACATGAACACC 4976
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrPgluGlyValPheThrGlyLeu 540
 Db 4977 CCGGGGCTTCCGCTGCGCAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCGCTC 5036
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 5037 ACTCATATAGATGCCCATTTCTATCCAGAACAGAGTGGGGGAGAACCTTCCCTTAC 5096
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
 Db 5097 CTGTAGCTACCAAGCACCGCTGTGCGTAGGGCTCAAGCCCTTCCCTCCATCGTGGGAC 5156
 Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 5157 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCCAACCTTCCATGCGGCCCAACACCCCTGCTA 5216

RESULT 5

US-08-440-519-9
 ; Sequence 9, Application US/08440519
 ; Patent No. 5712087
 ; GENERAL INFORMATION:
 ; APPLICANT: Houghton, Michael
 ; APPLICANT: Chco, Qui-Lim
 ; APPLICANT: Kuo, George
 ; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
 ; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Chiron Corporation
 ; STREET: P.O. Box 8097 (Int. Prop. R-440)
 ; CITY: Emeryville
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94662-8097
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: 12-MAY-1995
 ; APPLICATION NUMBER: US/08/440,519
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/910,760
 ; FILING DATE: 07-JUL-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Blackburn Esq., Robert P.
 ; REGISTRATION NUMBER: 30,447
 ; REFERENCE/DOCKET NUMBER: 0101.002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (510) 601-2702
 ; TELEFAX: (510) 655-3542
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9401 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 342..9374
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 366
 ; OTHER INFORMATION: /note= "this amino acid position
 ; can also be Arg."

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402

```

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also by Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-519-9
Alignment Scores:
Pred. No.: 0
Score: 3574.00
Length: 9401
Percent Similarity: 99.42%
Matches: 672
Best Local Similarity: 97.96%
Conservative: 10
Query Match: 98.78%
Mismatch: 4
Indels: 0
Gaps: 0
US-09-930-591-2 (1-686) x US-08-440-519-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3417 CTGGGGCCCATCAGCGGCTACGCCACGACACAGGGGCTCTTAGGTGCTAATACACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThra 40
DB 3477 AGCCTAACTGGCCGGGACAAAAACCAAGTGGAGGGTGAGGTCCAGATTGTCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 3537 GCCCAACCTTCCTGGCAACGTGTCATCATGGGGTGTCTGGACTGTCTACACGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValasp 80
DB 3597 GGAACGAGGACCATCGCGTCACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3657 CAAGACCTTGCGGCTGGCCCGCTCCGCAAGGTAGCGGCTCATACACCTGCATTCG 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaaspValIleProValArgArg 120
DB 3717 GGCTCTCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCTGGCGCGCG 3776

```

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer 140
DB 3777 GGTGATAGCAGGGGCGACCTGCTGTCGCCCGCCCATTTCTTACTTGAAGGCTCCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3837 GGGGGTCCGCTGTGTGTCGCCCGGGGCGACCGCTGGGCATATTTAGGGCCGGGTGTC 3896
QY 161 ThrArgGlyValAlaLysAlaValAppPheIleProValGluSerLeuGluThrThrMet 180
DB 3897 ACCCGTGGAGTGGCTAAGCGGGTGAATTTATCCCTGTGGAGAACCTAGAGACAACCATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
DB 3957 AGGTCCCGGTGTTCACGGATTAATCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLeuSerThrLysValProAlaAlaTyrAla 220
DB 4017 GCTCACCTCCATGCTCCACAGCGAGCGGCAAAAGCACCAAGGTCGCCGCTGCATATGCA 4076
QY 221 AlaGlyGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4077 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTCTGCAACACTGGGCTTTGGT 4136
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 4137 GCTTACATGTCGAAGGCTCATGGATCGATCCTTAACATCAGGACCGGGTGAGAACAT 4196
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaPheGlyCys 280
DB 4197 ACCACTGGACGCCCATCACTGATCCACCTAGCGCAAGTTCCTGCCAGCGGGGTGC 4256
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 4257 TCGGGGGGCGCTTATGACATATAATTTGTGACGAGTGCCTCCACGGATGCCACATCC 4316
QY 301 IleuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4317 ATCTTGGGCATCGGCATGCTTCATGACCAAGACGAGACTCGGGGGCGAGACTGGTGTG 4376
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4377 CTGCGCACCGCCACCCCTCGGGCTCGCTCAGTGTGCCCTCCACATCGAGGAGTT 4436
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 4437 GCTCTGTCCACCGCAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 4496
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 4497 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAAGGAAGTGCACGAACTCGCCGCA 4556
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4557 AAGTGTGTCATTTGGGCATCAATCGCTGGCCCTACTACCGCGTCTTGACGTGTCCGTC 4616
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 4617 ATCCCGACCGCGCGATGTTGCTCGTGGCAACCGATGCCCTCATGACCGGCTATACC 4676
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4677 GGCACCTTCGACTCGGTGATAGATGCAATACGTGTGTCCACGACAGCTCGATTCAGC 4736
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4737 CTTGACCCCTACCTTCACCAATTGACACAATCAGCTCCCGGATGTGTCTCCGCACT 4796
QY 461 GlnArgGlyArgThrGlyArgGlyProGlyIleTyrArgPheValAlaProGly 480
DB 4797 CAACGTCGGGCGAGACTGCGAGGGGAAGCCAGGCATCTACAGATTTGTGGACCGGGG 4856
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500

DB 4857 GAGCCCCCTCCGGCATGTTGCACTGCTCCCTCTCTGTGAGTGCTATGACGCGGCTGT 4916
QY 501 AlaTprTyrGluLeuThrThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4917 GCTTGATAGTACGTCAGCCCGCGAGACTACAGTAGGCTACGAGCGTACATGAACACC 4976
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
DB 4977 CCGGGGCTTCCCGTGTCCAGGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCGTC 5036
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 5037 ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 5096
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerThrAsp 580
DB 5097 CTGGTAGGCTACCAAGCCAGTGTGCGCTAGGCTCAAGCCCTCCCTCCCATCGTGGGAC 5156
QY 581 GlnMetTprLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5157 CAGATGTGGAAGTGTTCATTCGCTCAAGCCCTCCATGGGCCAACCCCTGCTA 5216
QY 601 TyrArgLeuGlyAlaValAlaGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 5217 TACAGACTGGGCGCTGTTCAGAAATGAAATCAACCTGACGCCACCCAGTCAACAAATACATC 5276
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTprValLeuValGlyGly 640
DB 5277 ATGACATCATGTGCGCGACCTGGAGTGTGTCAGGACCTGGGTGCTCTGTTGGCGGC 5336
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 5337 GTCTGTGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTGCTGTGTATAGTGGCGAG 5396
QY 661 IleValLeuSerGlyProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 5397 GTCTGTGCTGTCGGGAAGCCGCAATCATCTGACAGGAAGTCTCTCTACCGAGTTC 5456
QY 681 AspGluMetGluGluCys 686
DB 5457 GATGAGATGGAAGATGC 5474

RESULT 6
US-08-440-549-9
; Sequence 9, Application US/08440549
; Patent No. 6312889
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; TITLE OF INVENTION: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,549
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992

ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: G101.002
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9401 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 342..9374
FEATURE:
NAME/KEY: misc_feature
LOCATION: 366
OTHER INFORMATION: /note= "This amino acid position
can also be Arg."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
can also be Tyr."

NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
can also be Pro."
US-08-440-549-9
Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: Gaps: 4

US-09-930-591-2 (1-686) x US-08-440-549-9 (1-9401)

QY 1 MetalAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
:::|||||

Db 3417 CTGGCGCCATCAGCGGTAGCCAGCAGACAGAGGGCGCTCTAGGGTGCCATAATCAC 3476
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCCTAACTGCGCGGCAAAAACCAAGTAGTGGAGGTGAGTCCAGATTGTCAACTGCT 3536
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAAACTTCTGGCAACGTGATCAATGGGGTGTCTGGACTGTCTACCAACGGGCG 3596
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAAACGAGGACCAATCGCGTCAACCAAGGTCCTGTCTACAGATGTATACCAATGTAGAC 3656
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGCTGGCGGCTCCGCAAGTAGCGGCTCATTGACACCTGCACCTTGC 3716
Qy 101 GlySerSerAspLeuThrValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3717 GGTCTCTCGGACCTTTACCTGGTCACGAGGCAGCCGATGTCAATTCGCGTGGCGCGG 3776
Qy 121 GlyAspGlyArgGlySerLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3777 GGTGATAGCGGGCAGCCTGTCTGCCCGCGGCGACCGCGTGGCATAATTAGGGCGCGGTGTC 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGTCCTGTTGTGCCCGCGGGCACCGCGTGGCATAATTAGGGCGCGGTGTC 3896
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3897 ACCCGTGGAGTGGTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGGTCCCGGTGTTCAGGATACTCTCTCCACCACTAGTGCCCGACAGCTTCCAGGTG 4016
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 4017 GCTCACCTCCATGCTCCACAGGCGAGCGGCAAAAGCAACCAAGTCCCGGTGTCATATGCA 4076
Qy 221 AlaGlnGlyLysLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4077 GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTCTGCTGCAACACTGGGCTTTGGT 4136
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4137 GCTTACATGTCGAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAAT 4196
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4197 ACCACTGGACGCCCATCAGCTACTCCACCTAGCGCAAGTTCCTTGGCGACGGCGGGTGC 4256
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGGCGGCTTATGACATAATAATTGTGACGAGTGCCACTCCACGAGTGCACATCC 4316
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCATCGGCATGTCCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTGTG 4376
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4377 CTCGCCACCGCACCCCTCCGGGTCCTGCTACTGTGCCCCATCCCAACATCGAGGAGTT 4436
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4437 GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4497 AAGGGGGGAGACATCTCACTCTCTGTCAATCAAGAAGAAGTGTGGCAGCACTGGCGGCA 4556

Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4557 AAGCTGGTGGCATTTGGGCATCAATCCGTGGCTACTACCGGGTCTTTCAGCTGTCCGTC 4616
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4617 ATCCCGACGAGCGGCGATGTGTCTGCTGGCAACCGATGCCCTCATGCCGCTATACC 4676
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4677 GCGCACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTGCATTTCAGC 4736
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4737 CTTGACCTTACCTTCCACCATTCAGACATCACGCTCCCGCAGGATGCTGTCTCCGCACT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4797 CAACGTCCGGGCGAGACTGGCAGGGGGAAGCAGGACATCTACAGATTTGTGGCACCGGG 4856
Qy 481 GluArgProSerGlyWetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4857 GAGCGCCCTCCGGCATGTTTCGACTCTCGCTCTCTGTGAGTGTATGACGCGAGCTGT 4916
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGTGTAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGAGGGCGTCTTTACAGGCTTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGATGCCACTTTCATCCACAGCAAAAGCAGAGTGGGAGAACCTTCTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 5097 CTGGTAGCGTACCAAGCACCGTGTGCGCTAAGGCTCAAGCCCTCCATGGGCCAACCCCTG 5156
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAAGTGTTCGATTCGCTCAAGCCACCCCTCCATGGGCCAACCCCTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGACTGGGCGCTGTTTCAGAAATCAACCTTGAGCGACCCAGTCCACCAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5277 ATGACATGATGTCGGCGACCTGGAGGTGCTCAGCAGACCTGGGTGCTGTTGGCGGC 5336
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCTTGGCTGCTTGGCGCGTATTTCCTGTCAACAGGCTGCTGCTATATAGTGGCAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIleThrProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCTGCTTGTCCGGAAGCGGCAATATACCTGACGGGAAGTCTCTTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAAGAGTGC 5474

RESULT 7

US-08-823-895A-25
; Sequence 25, Application US/08823895A
; Patent No. 6433159
; GENERAL INFORMATION:
; APPLICANT: Kevin P. Anderson
; TITLE OF INVENTION: Compositions And Methods For
; TITL OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases

```

; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823.895A
; FILING DATE: March 17, 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/453,085
; FILING DATE: May 30, 1995
; APPLICATION NUMBER: 07/945,289
; FILING DATE: September 10, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 810-1454
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401
; TYPE: Nucleic
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: NO
; US-08-823-895A-25

Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservativeness: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-08-823-895A-25 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
Db 3417 CTGGCGCCCATCAGCGGTCAGCCAGCAGAGGGGCTCTTAGGGTGCATATCACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCCTAACTGCGCGGCAAAACCAAGTGGAGGTGAGTCCAGATTGTGCACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAAACTTCCTGGCAACGTGATCAATGGGGTGTGCTGGAGTGTCTACCAACGGGGC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAACGAGGACATCGCGTACCCAAAGGTCTGTATCCAGATGATATCAATGATAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGGTGCGCGCTCCGCAAGGTAGCGCTCAITGACACCTGCATCTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgAspArg 120
Db 3717 GGCTCCTCGGACCTTTACCTGTGTACGAGGACCCGATGTCATTCCTCCGTGCGCGGCG 3776
;
121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3777 GGTGATAGCAGGGGCGAGCTGCTGTCGCCCGGCCCATTTCTACTTGAAGAGTCTCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGGTCCGCTGTTGTCGCCCGCGGCGACGCCGTGGGCATATTTAGGGCCCGGTGTGC 3896
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
Db 3897 ACCGTGAGTGGCTAAGCGGTGACTTTATCCTGTGGAGAACCTTAGACACCAACATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGGTCCCGGTGTTCCAGGATAACTCTCTCCACAGTAGTGCCTCCAGAGTTCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 4017 GCTCACCTCCATGCTCCACAGCGAGCGGCAAGACCAAGGTCCCGGCTGCATATGCA 4076
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4077 GCTCAGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGT 4136
QY 241 AlaTyrMetSerLysAlaHisGlyLysLeuAspProIleAspProIleAspThrGlyVal 260
Db 4137 GCTTACATGTCCAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 4196
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys 280
Db 4197 ACCACTGGCAGCCCATCAGTACTCCACCTACGGCAGTCTTGGCGAGGGGGTGC 4256
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGCGGCTTATGACATAATAATTTGACAGTGCCTCCACGATCCACGATCCACATC 4316
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCATCGGCACCTGCTTGCACCAAGCAGACTGCGGGGGCGAGCTGGTGTG 4376
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4377 CTCGCCACCGCCACCCCTCGGGCTCCGTCACCTGTGCCCATCCCAACATCGAGAGTT 4436
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4437 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAGGCTATCCCTCCGAGTAAATC 4496
QY 361 LysGlyGlyValArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 4497 AAGGGGGGAGACATCTCATCTCTGTCATTCAAAGAGAGTGGAGCACTCGCGCA 4556
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4557 AAGCTGGTGGCATCAATGCCGTGGCTACTACCGCGGTCTGTAGCTGCTGCTC 4616
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4617 ATCCCGACAGCGGCGATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGTATACC 4676
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4677 GCGCATCTCGATGATAGACTGCAATACGTGTGTCCACCCAGACAGTGCATTTTCAGC 4736
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4737 CTGACCTTACCTTCCACATTGAGACATACGCTCCCGGAGATGCTGTCTCCGCACT 4796
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4797 CAACCTCGGGCAGGACTGGCAGGGGAGCAGGACATCTACAGATTTGTGGCACCGGG 4856
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500

```


Db 4857 GAGCGCCCTCCGCGATGTTCCGACTGCTCGTCTCTGAGTGCTATGACGAGGCTGT 4916
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGGTATGAGTCACGCGCGCGAGACTACAGTATAGGCTACGAGGTACATGAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTCCCGTGTCCAGGACCATCTTGAATTTGGAGGGCGTCTTTACAGGCCCTC 5036
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5097 CTGGTAGCTACCAAGCACCGTGTGCGTAGGGCTCAAGCCCTCCCCCATCGTGGAC 5156
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAAGTGTGTTGATTTGGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGCTGGCGCTGTTCAGATGAATCACCCTGACGACCACCGAGTCACCAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5277 ATGACATGATGTGCGCGACCTGGAGGTGCTCACGAGCACCTGGGTGCTCGTTGGCGGC 5336
Qy 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCTGTGCTGTGGCGCGGTATTGCTCTCAACAGGCTGCGTGTGTCATAGTGGCAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIlePheProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCTGTGTTCCGGAGCGCGCATCATCTGACAGGAGTCTCTTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAAGAGTGC 5474

RESULT 8

US-08-444-818-65
; Sequence 65, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisha A.

REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 6785 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..6785
US-08-444-818-65
Alignment Scores:
Pred. No.: 0 Length: 6785
Score: 3571.00 Matches: 671
Percent Similarity: 99.42% Conservativeness: 11
Best Local Similarity: 97.81% Mismatches: 4
Query Match: 98.70% Indels: 0
DB: 3 Gaps: 0
US-09-930-591-2 (1-686) x US-08-444-818-65 (1-6785)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1203 CTGGCGCCCATCACGGCTAGCCCCAGCAGCAAGGGGCCCTCTAGGCTGCATATCACC 1262
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1263 AGCCTTAACCTGGCGGGACAAAACCAAGTGGAGGCTGAGTCCAGATTGTGTCAACTGCT 1322
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db 1323 GCCCAAACTTCTCGCAACGTCATCATATGGGTGTCTGGACTGTCTACACCGGGGCC 1382
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1383 GGAACGAGGACCATCGCGTCACCCAAAGGTCCTGTCTATCCAGATGTATACCAATAGAC 1442
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1443 CAAGACCTTGTGGGTGGCCGCTCCGAGAGTAGCCGCTCATTCACACCTTGCCTTGC 1502
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 1503 GGCTCTCGGACCTTACTCTGTACAGGACGCGCGCATGTCTCCGTGCGCGGGGG 1562
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProLysSerTyrLeuLysGlySerSer 140
Db 1563 GGTGATAGCAGGGGAGGCTGTCTGCGCCCGGCCCATTTCTCTTGAAGGCTCTCTCG 1622
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 1623 GGGGTCTCGCTGTGTGCGCGCGGGCACGCGGTGGGCATATTTAGGGCCGCGGTGTC 1682
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1683 ACCCGTGGAGTGGCTTAAGCGGTGGACTTATCTCCCTGTGGAGAACCTAGAGAACCACTG 1742
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 1743 AGGTCCCGGTGTTTCACGGATAACTCTCTCCACAGTAGTGGCCCGCAGAGCTCCAGGTG 1802
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 1803 GCTCACCTCCATGCTCCACAGGAGCGGCAAAAGCAACCAAGGTCCCGGCTGCATATGCA 1862
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

Db 1863 GCTCAGGGCTATAGGTCCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 1922
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1923 GCTTACATGTCACAGGCTCATGGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAAATT 1982
Qy 261 ThrThrGlySerProIleThrTyrSerThrThrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 1983 ACCACTGGACGCCCATCACGTATCTCCACCTACCGCAAGTTCTTTGCGGACGGGGGTGC 2042
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2043 TCGGGGGCGCTTATGACATAATAATTGTGACGACTGCCACTCCACGGATGCCACATCC 2102
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2103 ATCTTGGGCATCGGCATCTCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTTGTG 2162
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2163 CTCGCCACCGCACCCCTCCGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGTT 2222
Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle 360
Db 2223 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 2282
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 2283 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAGAGTGGCGAGCTCGCGCA 2342
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2343 AAGCTGTGCGATTGGGCATCAATCGCGTGGCTTACTACCGCGCTCTTGACGTGTCGTC 2402
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2403 ATCCGACACGGCGGATGTTGTGCTGTGCAACCGATGCCCTCATGACCGGCTATACC 2462
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2463 GCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACCCAGACAGTCGATTTGAGC 2522
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2523 CTTGACCTTACCTTACCATTTGACCAATGACCAATCACGCTCCCGCAGATGCTGCTCCGCACT 2582
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2583 CAACGTGGGGCAGGACTGGCGAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 2642
Qy 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2643 GAGCGCCCTCCGGCATGTTCCGACTCGCTCGCTCTGTGAGTGCTATGACGAGGCTGT 2702
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2703 GCTTGGTATGAGCTTCACGCCCGCCGAGACTACAGTTAGGTTACGAGGTTACATGACACC 2762
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 2763 CCGGGGCTTCCCGTGTGCGAGGACATCTTGAATTTTGGGAGGGCGTCTTTACAGGCGCTC 2822
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2823 ACTCATATAGATGCCACTTCTTATCCAGCAAGCAGAGTGGGAGAACCTTCCTTAC 2882
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 2883 CTGGTAGCGTACCAAGCACCGCTGTGGCTAGGCTCAAGCCCTCCCTCCATCGTGGAC 2942
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2943 CAGATGTGGAAGTGTGTTGATTCGGCTCAAGGCCCACTTCCATGGGCGCAACACCCCTGCTA 3002

Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3003 TACAGACTGGGGCGCTGTTTCAGAAATACCTTACCTGACGCCACCCAGTCCCAAAATACATC 3062
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3063 ATGACATGCATGTCCGCCGACCTGGAGGTCTGCAGGACACCTGGTGTCTGTTGGGCGC 3122
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3123 GTTCCTGGCTGCTTTGGCGCGTATTGCTCTGTCACAGGCTCGTGGTCATAGTGGGCGAG 3182
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3183 GTCTGCTTGTCCGGAGACCGGCAATCATACTGCAGGGAAGTCCCTCTACCGAGAGTTC 3242
Qy 681 AspGluMetGluGluCys 686
Db 3243 GATGAGATGGAAGAGTGC 3260

RESULT 9

US-08-444-818-88
; Sequence 88, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 359-3876
; TELEFAX: (508) 359-3885
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8316
; US-08-444-818-88

Alignment Scores:
Pred. No.: 0
Score: 3570.00
Percent Similarity: 99.27%
Best Local Similarity: 97.81%
Length: 8316
Matches: 671
Conservative: 10
Mismatch: 5

Query Match:	98.67%	Indels:	0
DB:	3	Gaps:	0
US-09-930-591-2 (1-686) x US-08-444-818-88 (1-8316)			
Qy	1	MetAlaProileThrAlaTyAlaGlnGlnThrArgGlyLeuLeuGlyCysIleThr	20
Db	2734	CTGGCCCATCAGCGGTACGCCAGAGACAGAGGGCCCTCCAGGGTGCATATCACC	2793
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla	40
Db	2794	AGCCTAACTGGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTTGTCAACTGCT	2853
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyHisGlyAla	60
Db	2854	GCCAAACCTCTCTGCAACAGTGCATCAATGGGGTGTCTGCATCTCTACACGGGGCC	2913
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyThrAsnValAsp	80
Db	2914	GGAAACGAGGACCATCGCTACCCCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC	2973
Qy	81	GlnAspLeuValClyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	2974	CAAGACTTGTGGGCGGGCCCTCCGCAAGTAGTACCCCTCATTCAGACACCTGCATTCG	3033
Qy	101	GlySerSerAspLeuTyLeuValThrArgHisAlaAspValIleProValArgArg	120
Db	3034	GGCTCTCGGACTTTTACCTGTGTACGAGGACGCGGATGTCAATCCGCTGCGCGCGGG	3093
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyLeuLysGlySerSer	140
Db	3094	GGTGATAGAGGGGACGCTGTGTGCGCCCGGCCCATTTCTACTTGAAGGCTCCTCG	3153
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValAlaGlyIlePheArgAlaAlaValCys	160
Db	3154	GGGGTCCGCTGTGTGCGCCCGGGGACGCGGTGGGCATATTTAGGGCGCGGTGTGC	3213
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuLeuThrThrMet	180
Db	3214	ACCGGTGGAGTGGCTTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGACAAACATG	3273
Qy	181	ArgSerProValPheSerAspAsnSerProProAlaValProGlnSerTyArgGlnVal	200
Db	3274	AGTCCCGGTGTTCAGGATAACTCTCTCCACAGTAGTGCCCGCAGAGCTTCCAGGTG	3333
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyAla	220
Db	3334	GCTCACCTCATCTCCACAGCAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA	3393
Qy	221	AlaGlnGlyTyLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	3394	GCTCAGGGCTATAAGGTGTGTAGTACTCAACCCCTCTGTGTGTGCACACTGGGCTTTGT	3453
Qy	241	AlaTyMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	3454	GCATTATGTCCCAAGGCTCATGGATTCGATCTCTAACATCAGGACCGGGGTGAGAACAATT	3513
Qy	261	ThrThrGlySerProIleThrTySerThrTyGlyLysPheLeuAlaAspGlyLysCys	280
Db	3514	ACCACTGGAGCCCATCAGTACTCCACTACGCAAGTTCCTTCCGACGGCGGGTGC	3573
Qy	281	SerGlyGlyAlaTyAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	3574	TCGGGGGGCGGTATTGACATAATAATTTGTGACGAGTGCCATCCACGAGTGCACATCC	3633
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	3634	ATCTTGGGCATCGGCATGTCTTACCAACAGCAGAGACTGCGGGGCGGAGACTGGTTGTG	3693
Qy	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	3694	CTGCCACCGCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCAGAGAGTT	3753

RESULT 10


```

4775 CAACGTCGGGCGAGGACTGCGAGGGGAGCCAGGCGATCTACAGATTGTGGCAGCCGGG 4834
481  GUAAGProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
4835 GAGCGGCCCTCCCGCATGTTCGACTCGTCGCTCTGTGAGTGTATACAGCGCTGT 4894
501  AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
4895 GCTTGTATGAGCTACCGCCCGCGAGACTACAGTTAGGCTAGCAGCGTACATGAACACC 4954
521  ProGlyLeuProValCysGlnAspHisLeuGluPheTyrPheTyrGluValPheThrGlyLeu 540
4955 CCGCGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGCGCTCTTTACAGCCCTC 5014
541  ThrHisLeuAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
5015 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 5074
561  LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
5075 CTGGTAGCGTACCAAGCCAGCGTGTGGCTTAGGGCTCAAGCCCTCCCGCATCGTGGGAC 5134
581  GlnMetTyrPlyCysLeuLeuArgLeuLeuPheProThrLeuHisGlyProThrProLeuLeu 600
5135 CAGATGTGGAGGTGTGTATCGCCCTCAAGCCCAAGCCCTCATGGGCCAACCCCTGCTA 5194
601  TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
5195 TACAGACTGGCGCTGTTCAGATGAATCAACCTGACGCCACCCAGTACCAAAATACATC 5254
621  MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
5255 ATGACATGCAATGTCGCGCGACCTGGAGTGTCTACGAGCACCTGGGTGTCTGTGGCGGC 5314
641  ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
5315 GTCTGGCTCTTGGCGCGTATTCCTGTCAACAGGCTGCTGTATAGTGGCGAG 5374
661  IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
5375 GTGCTCTTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 5434
681  AspGluMetGluGluCys 686
5435 GATGAGATGGAGAGTGC 5452

```

```

RESULT 11
US-08-444-818-123/c
Sequence 123, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 123:
SEQUENCE CHARACTERISTICS:
LENGTH: 9185 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ANTI-SENSE: YES
US-08-444-818-123

Alignment Scores:
Pred. No.: 0 Length: 9185
Score: 3570.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5
Query Match: 98.67% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-123 (1-9185)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 5791 CTGGCGCCCATCACGGCGTACGCCAGCAGCAGCAGAGGGGCTCTCTAGGGTGCATATACACC 5732
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 5731 AGCCTAACTGGCGCGGACAAAACCAAGTGGAGGGTCCAGATTGTCTCAACTGCT 5672
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 5671 GCCCAAACTTCTGGCAACGTGCATCAATGGGTGTCTGACGTCTACACGGGCCC 5612
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 5611 GGAACGAGGACCATCGCGTCACCCAAAGGGTCTCTGTCAACAGATGTATACCAATGTAGAC 5552
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 5551 CAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCCTGCACITGC 5492
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 5491 GGCTCTCTGGACCTTTACTTGTTCAGAGGACGCGCGCATGTCATTCGCTGCGCGCGCGG 5432
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 5431 GGTGATAGCAGGGCAGCGCTGTGTGCGCCCGCGCCCACTTCTCTACTTGAAGAGGTCTCTCG 5372
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 5371 GGGGTCTCGCTGTGTGTGCGCGCGGCGACCGCTGGGCAATTTAGGCGCGCGGTGTGC 5312
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 5311 ACCCGTGGAGTGGCTAAGCGGTGGACTTATCTCTGTGGAGAACCTTAGAGACCAACCATG 5252
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrClnVal 200
Db 5251 AGGTCCCGGGTTCACGGATAACTCTCTCCACAGTAGTGGCCCCAGAGTCTCCAGGTG 5192
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 5191 GCTCACTCTCATGTCTCCACAGGAGCGGCAAGACCAACAGTCCCGGCTGCATATGCA 5132

```


QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrIleValAsp 80
DB 181 GGAACGAGGACCATCGCGTCACCAAGGCTCTGTATCATCAGATGATATACCAATGATAGAC 240
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 241 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTAGCGGATCATTCACACCCCTCACTTGC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGTCTCTCGGACCTTACCTGGTTCAGAGGACGCGGATGTCATCCCGTGGCCGCGCG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 361 GGTGATAGCAGGCGACCTGCTGCTCGCCCGGCCCATTTCTTACTTTGAAGGCTCTCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 421 GGGGGTCCGCTGTGTGTCGCCCGGGGACGCGGTGGGCATATTTAGGGCCGCGGTGTC 480
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 481 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGCAACCATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTCCCGCGTGTTCACGGATAACTCTCTCCACCATAGTGTGCCCGCCAGAGCTTCCAGGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 601 GCTCACTCTCATGCTCCACAGGAGCGGCAAGCCAGGTCCTCGGCTGCTATATGCA 660
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTCTGCTCAACACTGGGCTTTGGT 720
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCCAGGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 781 AACACTGGCAGCCCATCATCGTACTCCACCTACGGCAAGTTCCTTCCCGCAGCGGGTGC 840
QY 281 SerGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 841 TCGGGGGCGCTTATGACATAATAATTGTGAGAGTGGCCACTCCACGATGCCATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 901 ATCTTGGGCATTGGCATGTCTTGACCAAGCAGAGACTCGGGGGGAGACTGGTTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 961 CTGCGCACCCGCCCTCCGCGCTCGCTCACTGTGCCCATCCCAACATCGAGGAGTT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1021 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTATC 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCTCTCTCTCTCAAAAGGAAGTGGCAGCAACTCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1141 AAGCTGTCTGATGGGCATCAATCCGCTGCTTACTACCGCGGTCTTGCAGTGTCCGTC 1200
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCCGCCCATCGCGGATGTGTGCTGTGGCAACCATGCCCTCATGACCGGCTATACC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

DB 1261 GCGCACTTCGACTCGGTGATAGACTGCATATCGTGTCTACCCAGACAGTGCATTTCCAGC 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCCCTACCTTCACCATTCAGACATACGCTCCGCCAAGATGCTGTCTCCGCACT 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTCGGGCGAGACCTGGCAGGGGGAAGCGAGCATCTACAGATTTGTGGCCACCGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1441 GAGCGCCCTCCGCGCATGTCGACTCGCTCTCTGTGAGTGTATGACGACGAGCTGT 1500
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGGTATGAGTTCACGCCCGCGAGACTACAGTTAGGTACGAGCGGTATGAAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
DB 1561 CCGGGGCTTCCGCTGTGCGAGCAACCATCTTGAATTTTGGAGGGCGCTCTTTACAGGCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1621 ACTCATATAGATGCCACTTTCTATCCAGACAAAGAGAGTGGGAGAACCTTCCTTAC 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
DB 1681 CTGGTAGGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCCCATCGTGGGAC 1740
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1741 CAGATGTGGAAGTGTGTGATTGCGCTCAAGCCACCTCCATGGGCCCAACCCCTGCTA 1800
QY 601 TyrArgLeuGlyAlaValAlaGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 1801 TACAGACTGGGCGCTGTTCAGATGAAATACCTTCAGCGCACCCAGTCCCAATATACATC 1860
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
DB 1861 ATGACATGCATGTCCGCGAGCTCGGAGTGTCTACAGACACTGGGTGCTGTGGCGGC 1920
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 1921 GTCCCTGGCTGTCTTGGCGCGTATTGCTGTCAACAGGCTGCTGCTCATGTAGTGGCAGG 1980
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 1981 GTCTGTCTTGTCCGGAAGCGGCAATCATCTGACAGGAGTCTCTTACCGAGAGTTC 2040
QY 681 AspGluMetGluGluCys 686
DB 2041 GATGATGGAAGATGC 2058
RESULT 13
US-09-881-654-1
; Sequence 1, Application US/09881654
; Patent No. 6632601
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESKS, Laura
; APPLICANT: GEORGE-NASCIEMTO, Carlos
; APPLICANT: COLT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / PP17039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811

; PRIOR FILING DATE: 2001-04-02
 ; PRIOR APPLICATION NUMBER: 607280,867
 ; PRIOR FILING DATE: 2001-04-02
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2058
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: representative NS3/4a conformational antigen
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2058)
 ; US-09-881-654-1

Alignment Scores:
 Pred. No.: 0 Length: 2058
 Score: 3565.00 Matches: 671
 Percent Similarity: 99.13% Conservative: 9
 Best Local Similarity: 97.81% Mismatches: 6
 Query Match: 98.54% Indels: 0
 DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-09-881-654-1 (1-2058)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1 ATGGCGCCCATCAGCGGTACGCCAGCAGACAGAGGGGCCCTCTAGGGTGCAATACACC 60
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
 DB 61 AGCCTAATGTCGCGGACAAACCAAGATGGAGGGTCCAGATTGTGTCAACTGCT 120
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 DB 121 GCCCAACCTTCCTGGCAACGTGATCAATGGGGTGTCTGGACTGTCTACCAACGGGGCC 180
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 181 GGAAAGGAGGACCATCGCGTCAACCAAGGGTCTGTATCAGATGTATACCAATGTAGAC 240
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
 DB 241 CAAGACCTTGTGGGCTGGCCGCTCCCAAGAGTAGCGGATCTTGACACCTCGACTTGC 300
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 301 GGCTCCTCGGACCTTTACCTGGTCAAGAGGACGCGCGATGCTATTCCTGGCGCGCGGG 360
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 361 GGTATAGAGGGGACGCGTGTCTGCCCGCGGCCATTTCTACTTGAAGGCTCCTCG 420
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 421 GGGGGTCCGCTGTGTCCCGCGGGGCAAGCGGTGGGCATATTTAGGCGCGGGTGTGC 480
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGluThrThrMet 180
 DB 481 ACCCGTGGAGTGGGTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGAACCATG 540
 QY 181 ArgSerProValPheSerAspAsnSerProProAlaValProGlnSerTyrGlnVal 200
 DB 541 AGGTCCCGGCTGTCAGGATTAATCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTG 600
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaLysAla 220
 DB 601 GCTCACCTCCATGCTCCCAAGGAGGCGGCAAGACCAAGTCCCGGTGCGATATGCA 660
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAspProSerValAlaAlaThrMetGlyPheGly 240
 DB 661 GCTCAGGGCTATAAGGTGGTGTAGTACTCAACCCCTCTGTCTGCAACACTGGGCTTTGGT 720

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 721 GCTTACATGTCAGGCTCATGGGATCATCTTACATCAGAGCCGGGGTGAGAACAAIT 780
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
 DB 781 ACCACTGGCAGCCCATCATCGTACTCCACCTACGCAAGTTCCTTGGCGAGCGGGTGC 840
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 841 TCGGGGGCGCTTATGACATAATAATTTGTACGAGTGCCTACCCAGATGCCACATCC 900
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 901 ATCTTGGGCATTTGGCATGTCTTGACCAAGCAGAGACTGGGGGGCGAGACTGGTGTG 960
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 961 CTTCGCCACGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGAGGTT 1020
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 DB 1021 GCTCTGTCCACACCGGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTATC 1080
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 DB 1081 AAGGGGGGAGACATCTCTCTGTCTTCAAAAGAAAGTGCAGCAACTCGCCGCA 1140
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 1141 AAGTGTGTGGCATTTGGGCATCAATGCCGTGCCCTACTACCGCGTCTTGTAGTGTCCGTC 1200
 QY 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 1201 ATCCCGCCCATCGGCGATGTTGTCGTGTGGCAACCGATGCCCTCATCAGCGGTATACG 1260
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 1261 GCGCACTTCGATCGGTGATAGCTGCAATACGTGTGTACCCAGAGATGCTGTCTCCGCACT 1320
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 1321 CTTGACCTCATCTTCCACCATTTGAGACATCACGCTCCCCAAGATGCTGTCTCCGCACT 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 DB 1381 CAACGTCGGGCGAGGACTGGCAGGGGGAAGCAGGCGCATCTACAGATTTGTGGCACCGGG 1440
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 DB 1441 GAGCGCCCTTCGGCATGTTTCGATCTGCTCTCTGTGAGTGTCTATGAGCGAGGTGT 1500
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 1501 GCTTGGTATGAGTCAACGCCCGCGAGACTACAGTAGGCTTACGAGCGTATCAACACACC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 DB 1561 CCGGGGCTTCGGGTGTCAGGACCATCTTGAATTTGGGAGGGCGCTTTTACAGGCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 DB 1621 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 1680
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580
 DB 1681 CTGTGTAGGCTTACCAAGCCACCGCTGTGGCTAGGCTCAAGCCCTCCCTCCATCGTGGAC 1740
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 1741 CAGATGTGGAAGTGTGTGATTTCGCTCAAGCCCTCCATCGGCGCAACACACCTCTCTA 1800

QY 601 TtArgLeuGlyAlaValGlnGlnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TACAGACTGGCGCTCTTCAAGATGAATCAACCTGACGCCACCCAGTCACCAATACATC 1860
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATACATGATGTCGGCCGACCTGGAGTCTGTCACGAGCACCTGGGTGCTCGTTGGCGGC 1920
QY 641 ValLeuAlaLeuAlaValThrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTCTGCTCTTGGCGCGCTATTGCTGTCACAGAGGTGCTGCTAGTGGGCAGG 1980
QY 661 IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 GTGCTCTTGGCGCGAGCGCGCAATCATACCTGACAGGAGTCTCTTACCGAGAGTTC 2040
QY 681 AspGluMetGluGluCys 686
Db 2041 GATGAGATGAAGAGTGC 2058

RESULT 14

US-08-444-818-137
; Sequence 137, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Ruter, William J.
; TITLE OF INVENTION: NANB Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/444,818
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8987 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8985
US-08-444-818-137

Alignment Scores:
Pred. No.: 0
Score: 3565.00
Percent Similarity: 99.27%
Best Local Similarity: 97.81%
Query Match: 98.54%
Length: 8987
Matches: 671
Conservative: 10
Mismatch: 5
Indels: 0

DB: 3 Gaps: 0
US-09-930-591-2 (1-686) x US-08-444-818-137 (1-8987)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3076 CTGGCGCCCATACAGCGGTACGCCAGCAGCAAGAGGGGCTCTCTAGGTGATATACACC 3135
QY 21 SerLeuThrGlyArgAspLeuValGlnGlyGluValGlnIleValSerThrAla 40
Db 3136 AGCTTAATCTGGCCGGGCAAAAAACAGTGGAGGTGAGGTCCAGATTGTGTCACTGCT 3195
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60
Db 3196 GCCCAAACTTCTGGCAACGTGCATCAATGGGTGTGTCTGGAGTGTCTACACGGGGCC 3255
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3256 GGACGAGGACCATCGCTACCAAGGGTCTCTCATCCAGATGTATACCAATGTAGAC 3315
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3316 CAGACCTTGTGGGTGGCTGGCTCCGCAAGGTAGCGCTCATTGACACCTCGACCTTGC 3375
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 3376 GGCTCTCGACCTTTACTGTGTACGAGGACGCGCATTCATTCCGGTGGCGGCGG 3435
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3436 GGTGATAGCAGGGGAGCTGTGTGCGCCCGGCCATTTCTACTTGAAGAGGTCTCTCG 3495
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3496 GGGGTCTCGTGTGTGCCCCGCGGCGCACGCGTGGGCATATTTAGGGCGCGGTGTGC 3555
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3556 ACCCGTGGAGTGGCTAAGCGGTGACCTTATCCCTGTGGAGAACCTAGACACACCATG 3615
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3615 AGGTCCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGTGCCCGCAGAGCTTCCAG 3675
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaThrMetGlyPheGly 240
Db 3676 GGTCACTCCATGCTCCACAGCAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA 3735
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 3736 GGTCAAGGCTAAGGTGCTAGTACTCAACCTCTGTGTGCAACACTGGGCTTTGGT 3795
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 3796 GCTTACATGTCCAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 3855
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 3856 ACCACTGGCAGCCCATCATCAGTACTCCACTACGGCAAGTTCCTTGCCGACGGCGGTGC 3915
QY 281 SerGlyGlyValAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 3916 TCGGGGGGCGCTTATGACATATAATTTGTGACGAGTGCCTCCACGATGCCACATCC 3975
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 3976 ATCTTGGGATCGGCACCTGCTCTTCAAGCAGAGACTGCGGGGGCGGAGACTGGTTGTG 4035
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal 340
Db 4036 CTCGCCACCGCCACCCCTCCGGGCTCCGTCTCTGTGCCCCCATCCCAACATCCAGGAGGT 4095
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360

Db 4096 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGTATATCCCTTCGAGTAATC 4155
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAala 380
Db 4156 AAGGGGGGGGAGACATCTCATCTTCTGCTATCAAGAGAGAGTGGCAGCACTCGCGCA 4215
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4216 AAGCTGTGTCGATTTGGGCATCAATGCGGTGCTACTACCGCGCTTTGACGTTCGTC 4275
QY 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4276 ATCCGACACCGCGGATGTTGTCTGCTGCGTGGCAACCGATCCCTCATGACCGCTATACC 4335
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4336 GCGGACITCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCGATTTCAGC 4395
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4396 CTTGACCTTACCTTACCATTTGACCAATGACCAATCACGCTCCCGAGGATGCTCTCCGCACT 4455
QY 461 GlnArgArgGlyArgThrGlyArgGlyValProGlyIleTyrArgPheValAlaProGly 480
Db 4456 CAACGTGGGGCAGACTGGCAGGGGAGCCAGGCATCAACAGATTGTGGCACCAGGG 4515
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4516 GAGCGCCCTCCCGCATGTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4575
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4576 GTTTGGTATGAGCTCAACCGCCGCGAGCTACAGTTAGGCTAGCAGCTACATGAACACC 4635
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4636 CCGGGCTTCCCGTGTCCAGGACCATCTTGAATTTGGAGGGCGCTCTTACAGGCTC 4695
QY 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 4696 ACTCATATAGATGCCACTTCTTATCCAGCAAGCAGAGTGGGAGAACCTTCTCTTAC 4755
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 4756 CTGGTAGCTACCAAGCAGCCGTGCTGCTAGGCTCAAGCCCTCCCTCCATCGTGGAC 4815
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 4816 CAGATGTGGAGTGTGTTGATTGCTCCTCAAGCCACCCCTCCATGGGCAACACCCCTGCTA 4875
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 4876 TACAGCTGGCGCTGTTCAATGAATGAATCACCCTCAGCCACCCAGTACCAAAATACATC 4935
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 4936 ATGACATGCTATGCGGCGACCTGGAGTGTGTCAACGACACCTGGTGTGTGGCGC 4995
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 4996 GTCTGTGCTGCTTGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5055
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5056 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5115
QY 681 AspGluMetGluGluCys 686
Db 5116 GATGATGGAAGAGTGC 5133

RESULT 15

US-08-444-818-53

/ Sequence 53, Application US/08444818
/ Patent No. 6150087
/ GENERAL INFORMATION:
/ APPLICANT: Chien, David Y.
/ APPLICANT: Rutter, William J.
/ TITLE OF INVENTION: NANBV Diagnostics and Vaccines
/ NUMBER OF SEQUENCES: 777
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Chiron Corporation
/ STREET: 4560 Horton Street
/ CITY: Emeryville
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94608-2916
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/444,818
/ FILING DATE:
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/403,590
/ FILING DATE: 14-MAR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Harbin, Alisa A.
/ REGISTRATION NUMBER: 33,895
/ REFERENCE/DOCKET NUMBER: 0110.002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (508)359-3876
/ TELEFAX: (508)359-3885
/ INFORMATION FOR SEQ ID NO: 53:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5360 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..5360
/ US-08-444-818-53

Alignment Scores:
Pred. No.: 0 Length: 5360
Score: 3555.00 Matches: 668
Percent Similarity: 99.42% Conservative: 13
Best Local Similarity: 97.52% Mismatches: 4
Query Match: 98.26% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-53 (1-5360)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 930 CTGGCGCCCATCATCGCGGTACGCCACACAGGGGGCTCTAGGTGCATAATCACC 989
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 990 AGCTTAATCTGCGCGGCAAAACCAAGTGGAGGTGAGTCCAGATTGTGTCACTGCT 1049
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1050 GCCCAACCTTCTGGCAACGTGCAATCAATGGGGTGTGTGGACTGTCTACCGGGGCC 1109
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1110 GGAACGAGGACCATCGCGTCACCCAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1169
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100

Db 1170 CAAGACCTTGTGGGCTGGCCCGCTCGGCAAGGTAGCGCGCTCATTGACACCCCTGCACATTGC 1229
Qy GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 1230 GGTCTCTCGGACCTTTACCTGTGTACAGGCAAGCGCGATGTATTCCCGTGGCGCGGG 1289
Qy 121 GlyAspGlyArgCysLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 1290 GGTGATAGCAGGGGAGCTGTCTGTGCGCCCGGCGCCATTTCTACTTGAAGGCTCTCTCG 1349
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaValCys 160
Db 1350 GGGGGTCCGCTGTGTGTGCGCGGCGGACGCGGTGGGCATATTTAGGCGCGGTGTGC 1409
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1410 ACCCGTGGAGTGGCTAAGCGGTGGATTTATCCCTGTGGAGAACCTAGAGCAACCCATG 1469
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 1470 AGGTCCCGGTGTTCAGGATTAATCTCTCCACAGTAGTAGTCCCGCAGAGCTTCCAGGTG 1529
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaValAla 220
Db 1530 GCTCACCTCCATGCTCCACAGCAGCGGCAAGCAAGCAAGCTCCCGCTGCATATGCA 1589
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 1590 GCTCAGGCTATAGGTGTGTAGTACTCAACCCCTCTGTGTGTGCAACATGGGCTTTGGT 1649
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1650 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 1709
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyPheLeuAlaAspGlyGlyCys 280
Db 1710 ACCACTGGCAGCCCATCAGTACTCCACTCAGCAAGTTCTTCCCGACGCGGTGC 1769
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 1770 TCGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCACATCCACGGATGCCATCC 1829
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 1830 ATCTTGGGCATCGGCATGCTCTTGACCAAGAGAGACTCGGGGGGAGACTGGTGTG 1889
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 1890 CTGCCACCGCCACCCCTCCGGGCTCGCTCACTGTGCCCATCCACATCCGAGAGGT 1949
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1950 GCTCTGCCACACCGGAGAGATCCCTTTTAAGGCAAGGCTATCCCGCTCGAAGTAATC 2009
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2010 AAGGGGGGAGACATCATCTTCTGTCAATTAAGGAAGAGTCCGACGAACTCGCGCA 2069
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2070 AAGTGTGTGGATCGCATCAAGTCCGCTGCTTACTACCGCGCTCTTGAGCTGTCCGTC 2129
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2130 ATCCCGACACGCGCGATGTGTGTGTGTCGCAACCCATGCCCTCATGACCGGCTATACC 2189
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2190 GCGACCTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCACAGACAGTCAATTCAGC 2249
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2250 CTTGACCCCTACCTTCACCATTTGAGACAAATCAGCTCCCGCAGGATGTGTCTCCCGCACT 2309

Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2310 CAACGTGGGGCAGACTGGCAGGGGAGCCAGGCACTTACAGATTGTGGCACCAGGG 2369
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2370 GAGCGCCCTCCGGCATGTTCCGACTCGTCCGTCTCTGTGAGTGCTATGACGAGGCTGT 2429
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2430 GCTTGGTATAGCTACGCCCGCGAGACTACAGTTAGGTACGAGGTACATGAACACC 2489
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 2490 CCGGGCTTCCCGTGTGCCAGACCATCTTCAATTTTGGAGGGCGCTTTTACAGGCGTC 2549
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProThr 560
Db 2550 ACTCATATAGATGCCCATTTCTATCCAGCAAGCAGAGTGGGAGAACCTTCTTCTTAC 2609
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 2610 CTGGTAGGTACCAAGCACCGCTGTGGCTAGGCTCAAGCCCTCCCGCATCGTGGAC 2669
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2670 CAGATGTGGAAGTGTGTTGATTGGCTCAAGCCCATCTCCATGGGCCAACACCCCTGCTA 2729
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 2730 TACAGACTGGGCGCTGTTCAGATGAATCACTCAGCGCACCCAGTCAACCAATACATC 2789
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640
Db 2790 ATGACATGCATGTGGCGGACCTGGAGGTCTGCACGACACCTGGTGTCTCGTTGGCGC 2849
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 2850 GTCTGTCTGTCTTGGCGCGTATTGCTGTCAACAGGCTCGTGTGTCTATAGTGGCAGG 2909
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 2910 GTCTGTCTGTCTGGGAGCGCGCAATCATCTGACAGGGAAGTCTCTCCTCAGAGTTC 2969
Qy 681 AspGluMetGluGlu 685
Db 2970 GATGAGTGAAGAG 2984

Search completed: September 16, 2004, 00:45:35
Job time : 330 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 15, 2004, 15:46:57 ; Search time 712 Seconds

(without alignments)
4093.067 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618

Sequence: 1 MAPITAYAAQTRGLGCIIT.....PALIPDREVLVREFDEMEEC 686

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model
-Q=/CGN2_1/USPTO.spool/US0930591/runat_15092004.164645.7086/app.query.fasta_1.839
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=ring -XINMATCH=0.1 -LOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0930591 @CGN 1.1 470 @runat_15092004.164645.7086 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002s:.*
7: Geneseqn2003as:.*
8: Geneseqn2003bs:.*
9: Geneseqn2003cs:.*
10: Geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3618	100.0	2061	6 AAD34500	Aad34500 Hepatitis
2	3618	100.0	2061	6 AAD31767	Aad31767 Hepatitis
3	3618	100.0	2061	9 AAD60868	Aad60868 Hepatitis
4	3574	98.8	7310	1 AAN92106	Aan92106 Combined
5	3574	98.8	7310	1 AAN90336	Aan90336 Composite
6	3574	98.8	7310	2 AAG98221	Aag98221 Hepatitis
7	3574	98.8	8316	3 AAA75296	Aaa75296 cDNA sequ
8	3574	98.8	9133	2 AAZ07656	Aaz07656 Nucleotid

9	3574	98.8	9185	2	AAQ10566	Aaq10566 Hepatitis
10	3574	98.8	9185	3	AAA75297	Aaa75297 Sense str
11	3574	98.8	9401	2	AAT12710	Aat12710 Hepatitis
12	3574	98.8	9401	2	AAT12710	Aat12710 HCV polyp
13	3574	98.8	9401	2	AAV09989	Aav09989 HCV polyp
14	3574	98.8	9401	6	AAV09989	Aav09989 HCV polyp
15	3571	98.7	6905	1	AAN92103	Aan92103 Combined
16	3571	98.7	9185	2	AAQ26737	Aaq26737 Nucleotid
17	3570	98.7	9185	2	AAQ05956	Aaq05956 Sense str
18	3568	98.6	6299	4	AAQ05956	Aaq05956 HCV NS3/4A
19	3568	98.6	9185	2	AAQ05956	Aaq05956 Hepatitis
20	3567	98.6	8316	2	AAQ05955	Aaq05955 Hepatitis
21	3565	98.5	2058	6	ABK15344	Abk15344 Hepatitis
22	3565	98.5	2058	6	AAD29795	Aad29795 HCV-1 NS3
23	3565	98.5	2058	7	ABX14410	Abx14410 DNA encod
24	3565	98.5	2058	9	ADC06768	Adc06768 HCV mutan
25	3565	98.5	9400	2	AAQ21744	Aaq21744 Compiled
26	3555	98.3	5360	1	AAN90327	Aan90327 Hepatitis
27	3552	98.2	5300	1	AAN92097	Aan92097 Hepatitis
28	3551	98.1	9646	2	AAV59361	Aav59361 Hepatitis
29	3551	98.1	9646	6	ABK87285	Abk87285 cDNAenco
30	3551	98.1	9646	7	ACA62466	Aca62466 HCV H77 C
31	3551	98.1	12980	2	AAV59364	Aav59364 Hepatitis
32	3551	98.1	12980	6	ABK87286	Abk87286 Hepatitis
33	3551	98.1	12980	7	ACA62469	Aca62469 DNA encod
34	3551	98.1	16622	3	AAZ36212	Aaz36212 Nucleotid
35	3550	98.1	9502	2	AAQ74770	Aaq74770 Hepatitis
36	3544	98.0	9518	5	AAD03778	Aad03778 Hepatitis
37	3544	98.0	9599	2	AAZ24833	Aaz24833 Infectiou
38	3544	98.0	9599	2	AAZ24832	Aaz24832 Infectiou
39	3544	98.0	9599	4	AAZ24832	Aaz24832 Infectiou
40	3544	98.0	9611	5	AAZ24832	Aaz24832 Infectiou
41	3544	98.0	9611	5	AAZ24832	Aaz24832 Infectiou
42	3544	98.0	9611	5	AAZ24832	Aaz24832 Infectiou
43	3544	98.0	9611	5	AAZ24832	Aaz24832 Infectiou
44	3544	98.0	10803	7	ABX10617	Abx10617 MKO-Z nuc
45	3543	97.9	9379	2	AAQ36209	Aaq36209 Composite

ALIGNMENTS

RESULT 1
AAD34500
ID AAD34500 standard; DNA; 2061 BP.
XX
AC AAD34500;
XX
DT 16-JUL-2002 (first entry)
XX
DE Hepatitis C virus NS3/4A protein encoding DNA.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; gene; ds.
XX
OS Hepatitis C virus.
XX
PH Key Location/Qualifiers
FT CDS 1..2061
FT /*tag= a
FT /product= "HCV NS3/4A protein"
XX
PN WO200214362-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001774.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX
PA (TRIP-) TRIPEP AB.
XX

QY	581	GlnMetTrpLysCysLeuIleA	gLeuLysProThrLeuHisgLyProThrProLeuLeu	600
Db	1741	CAGATGTGGAGTGTTCATCCGTCT	CAAGCCACCCCTCCATGGGCCAACACCTCTGCTA	1800
QY	601	TyrArgLeuGlyAlaValAla	GlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	1801	TATAGACTGGGCGCTGTCCAGATGAAGT	CACCTGACGACCCAGTCACCAAGTATATC	1860
QY	621	MetThrCysMetSerAlaAsp	LeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	1861	ATGACATGTATGTCCGGCTGACCTGGAGT	CGTCACGAGTACCTGGGTGCTCGTTGGCGGC	1920
QY	641	ValLeuAlaAlaLeuAlaAla	TyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	1921	GTTCCTGCTCTTTGGCGCGGTATTGCTAT	CCATCCACAGCTCGGTGGTCATAGTAGGAGG	1980
QY	661	IleValLeuSerGlyLysPro	AlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	1981	ATTGTCTTGTCCGAAAGCCGCAAT	CATATCCCGACAGGGAAGTCTCTACCGGAGTTC	2040
QY	681	AspGluMetGluGluCys	586	
Db	2041	GATGAATGGAAGAGTGC	2058	
RESULT 2				
AAD31767				
ID	AAD31767 standard; DNA; 2061 BP.			
XX	AAD31767;			
AC				
XX	18-JUN-2002 (first entry)			
XX	Hepatitis C virus (HCV) NS3/4A DNA coding region.			
DE	Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;			
XX	cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;			
KW	ds.			
XX				
OS	Hepatitis C virus.			
XX				
XX				
Key	Location/Qualifiers			
PH	1..2061			
FT	/*tag= a			
FT	/product= "HCV NS3/4A protein"			
XX	WO200213855-A2.			
PN				
XX	21-FEB-2002.			
PD				
XX				
XX	15-AUG-2001; 2001WO-IB001808.			
XX				
XX	17-AUG-2000; 2000US-0225767P.			
PR	29-AUG-2000; 2000US-0229175P.			
PR	03-NOV-2000; 2000US-00705547.			
XX				
XX	(TRIP-) TRIPEP AB.			
PA				
XX				
XX	Sallberg M, Hultgren C;			
PI				
XX	WPI; 2002-241837/29.			
DR	P-PSDB; AAE19900.			
XX				
XX	Vaccine compositions for treating and preventing disease, preferably			
PT	hepatitis C virus infection, comprises ribavirin and antigen that has			
PT	epitope present in hepatitis C virus.			
XX				
PS	Claim 1; Page 94-95; 120pp; English.			
XX				
CC	The invention relates to a composition comprising ribavirin and an			
CC	antigen preferably non structural 3 protein (NS3)/4A fragment of			
CC	hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV			
CC	sequence. The composition is useful for enhancing an immune response to a			
CC	hepatitis C antigen in humans, domestic, sport or pet species and as			
CC				

CC	XX	SQ	Alignment Scores:	2.44e-245	Length:	2061
CC	XX	SQ	Pred. No.:	3618.00	Matches:	696
CC	XX	SQ	Percent Similarity:	100.00%	Conservative:	0
CC	XX	SQ	Best Local Similarity:	100.00%	Mismatches:	0
CC	XX	SQ	Query Match:	100.00%	Indels:	0
CC	XX	SQ	DB:	6	Gaps:	0
CC	XX	SQ	Sequence	2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;		
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	SQ	Alignment Scores:			
CC	XX	SQ	Pred. No.:			
CC	XX	SQ	Percent Similarity:			
CC	XX	SQ	Best Local Similarity:			
CC	XX	SQ	Query Match:			
CC	XX	SQ	DB:			
CC	XX	SQ	Sequence			
CC	XX	S				

Db 841 TCAGGGGTGCTATGACATAATATTTGTACGAGTGCCTCCAGGATGCACATCC 900
 QY 301 ILeuLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 901 ATCTTGGGCATGGACATGCTCTTGACCAACAGAGACCGGGGGGAGACTGACTGTG 960
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 961 CTCGCCACCGCTACCCCTCCGGGCTCGTCACTGTGCCCTCCATCAATCAACGAGAGGTT 1020
 QY 341 AlaLeuSerThrThrGlyIleProPheThrGlyValAlaIleProLeuGluAlaIle 360
 Db 1021 GCTCTGTCCATACCGGAGAGATCCCTTTATGGCAAGGCTATTCCTCTTGAGCAATT 1080
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 Db 1081 AAGGGGGGAGACATCTCATCTCTGCCACTCAAGAAGAGTGCAGAGCTCGCCGCA 1140
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaIleThrValArgGlyLeuAspValSerVal 400
 Db 1141 AAACGTGTCGGGTGGGGGTCAATGCGGTGCTTACTACCGCGGCTTGATGTGTCGTC 1200
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 1201 ATCCGACCACTGGTGAGTGTCTGCTGGCAACTGACGCCCTCATGACCGGCTTACC 1260
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 1261 GCGCACTTCGATTCGGTGATAGACTCAACAGCTGTGTCAACAGAGAGTGCAGCTTCAGC 1320
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 1321 CTTGACCTTACTTCCATTTGACAGCAATCAAGCTTCCAGAGTGTCTGTCTCCGTACT 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyProGlyIleThrArgPheValAlaProGly 480
 Db 1381 CAACGTGCGGGTAGGACTGGCAGAGGAAAGCAGGATCTACAGATTTGTGGCACCAGGG 1440
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysThrAspAlaGlyCys 500
 Db 1441 GAGCGTCTCTTGGCATGTTTGAATCTGCTGCTCTGCGAGTGTCTGAGACGGGGTGT 1500
 QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 1501 GCTTGTGTATGACTACCGCCCGGACACACAGTTAGGCTACGAGCATACATGAACACC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 1561 CCGGGACTTCCGTTGTCGCAAGACCATCTTGAATTTTGGAGGGCGGTCTTTACGGTCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluLeuProTyr 560
 Db 1621 ACCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 1681 CTGTAGCGTACCAAGCACCGTGTGGGTAGAGCTCAAGCCCTCCCGCGTGTGGGAC 1740
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGGAGTGTGTATCCGCTCAAGCCACACCTCCATGGGCCAACACCTCTGCTA 1800
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 1801 TATAGACTGGCGGTGTGCAAGTGAAGTCACTGACGACCCAGTCAACCAAGTATATC 1860
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 1861 ATGACATGTATGTCGGTGTGCTGGAGTGTGTACGAGTACCTGGGTGTCTGTGGCGC 1920
 QY 641 ValLeuAlaLeuAlaIleThrCysLeuSerThrGlyCysValValIleValGlyArg 660

Db 1921 GTTGTGGTGTCTTTGGCCGCGTATTGCTCTATCCACAGGCTCGTGTCTATAGTAGG 1980
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 1981 ATTGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGAAGTCTCTACCGGGAGTTC 2040
 QY 681 AspGluMetGluGluCys 686
 Db 2041 GATGAATGGAAGAGTGC 2058

RESULT 3
 AAD60868
 ID AAD60868 standard; DNA; 2061 BP.
 XX AAD60868;
 DT 15-JAN-2004 (first entry)
 XX Hepatitis C virus NS3/4A DNA.
 XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
 KW virucide; ds.
 XX Hepatitis C virus.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..2061
 FT /*tag= a
 FT /*product= "Hepatitis C virus protein"
 XX US2002136740-A1.
 XX 26-SEP-2002.
 PF 15-AUG-2001; 2001US-00929955.
 XX 17-AUG-2000; 2000US-0225767P.
 PR 29-AUG-2000; 2000US-0229175P.
 XX (SALL/) SALLBERG M.
 PA (HULT/) HULTGREN C.
 XX Sallberg M, Hultgren C;
 DR WPI; 2003-764978/72.
 DR P-PSDB; ABW00351.
 XX Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 XX Claim 1; Page 60-61; Opp; English.
 CC The invention relates to a composition comprising ribavirin and an
 CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
 CC is useful in enhancing the immune response to a hepatitis C antigen where
 CC the composition is delivered to an animal identified as requiring an
 CC enhanced immune response. The vaccine is useful in the treatment and
 CC prevention of hepatitis C infection. The present sequence is Hepatitis C
 CC virus NS3/4A DNA
 XX
 SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.44e-245 Length: 2061
 Score: 3618.00 Matches: 686
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 9
 US-09-930-591-2 (1-686) x AAD60868 (1-2061)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1 ATGGCGCTATCAGCGCTATGCCAGCAGACAAAGGGGCTTTTGGATGCATAATCAC 60
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 61 AGCTTGACCGCGCGGACAAACACAGGTGGAGGGTGAGGTTTCAGATCGTGCACTGCT 120
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 121 GCCCAGACTTCTTGGCAACTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCC 180
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 181 GGAACAGGACCAATGGCTCACCTAGGTCCTGTATCCAGATGTACACCAATGTGGAC 240
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTCTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCATCTGC 300
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 301 GGCTCCCTCGGACCTTTACCTGGTTCAGGACAGCCGATGTCAITCTGTGCCCGCACGG 360
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 361 GGTGATGGCAGGGCAGCCTGCTTCCGCCCGGCTATCTCTTACTTTGAAAGGCTCCTCG 420
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGAGGCCCTCTGTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGGGTATGC 480
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 481 ACCCGTGGATGGCTAAGCGGTGGACTTCATCCCCGTAGAGACTAGAGACAACCATG 540
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 541 AGTCCCGGTGTCTCAGACAATCTCTCCCGACAGCAGTCCCGCAGAGCTACCAAGTG 600
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 601 GCCCACTGTATGCTCCACCGCGGCGGGTAAGACACCAAGTCCCGCGCGCATACGCA 660
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 661 GCTCAGGGCTACAAGGTGCTGGTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGT 720
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCCAAAGGCCCATGGGATGATCCTAACATCAGGACTGGGGTGAGACAATT 780
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
Db 781 ACTACTGGCAGCCGATCAGTATTCACCTAGGCAAGTTCTTGGCCGACGGGGTGT 840
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 841 TCAGGGGGTGTATGACATAATAATTTGTACAGAGTGCACCTCCACGGATGCAACATCC 900
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTTGGCATTTGCACTGTCTTGACCAAGCAGAGACCGGGGGGGGAGACTGTGTG 960
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 961 CTCGCCACCGTACCCCTCCGGGCTCCGTCCTACTGTGCCCATCTCTAACATCGAGGGT 1020
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATCTCCCTTGAAGCAAT 1080
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380

Db 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAAGAGTGGAGAGCTGCGCGCA 1140
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 1141 AAACCTGGTGGCGTTCATGCGCTGCTTACTACCGCGGCTTGATGTGTCCGTC 1200
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCCGACCATGTGTGACGTGTGTGTGGCAACTGACGCCCTCATGACCGGTTTACC 1260
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1261 GGCAGACTTCGATTCGATAGACTGCAACACGCTGTGTCAACCCAGACAGTGCATTCAGC 1320
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCTTACCTTACCATTGAGACAATCAGCTTCCCCAGGATGCTGTCTCCGTA 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 1381 CACGCTCGGGTAGACTGGCAGAGGAAAGCAGCACTCTACAGATTTGTGGCACCGGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1441 GAGCGTCTCTTGGCATGTTTACTCTGTCTGTCTGCGAGTGCTATGACGCGGGTGT 1500
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1501 GCTTGTATGAGCTTACGCCCGCGAGACACAGCTTAGGCTACGAGCATACATGAACAC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 1561 CCGGACCTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTTTTACGGGTCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1621 ACCCACATAGACGCCCACTTCTATCCAGACAAGCAGAGTGGGGAACACCTTCCCTAT 1680
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 1681 CTGGTAGCGTACCAAGCCACCGTGTGCGTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAAGTGTGTATCGCTCAAGGCCACCCCTCCATGGGCCAACACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TATAGACTGGCGCTGTCCAGATGCAAGTCAACCTGACGACCCAGTCAACCAAGTATATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGTATGTCCGCTGACCTGGAGTGTGTGACAGTACTCTGGTGTCTGTTGGCGGC 1920
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGGCTGTCTTGGCGCGGTATTCCTATCCACAGGCTGCGTGTCTATAGTAGTAGG 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 ATTGTCTTGTCCGAAAGCCGGCAATCATACCCAGACGGAAGTCTCTACCGGGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAATGGAAGAGTGC 2058
RESULT 4
AAN92106
ID AAN92106 standard; DNA; 7310 BP.
XX
AC AAN92106;
XX

DT 25-MAR-2003 (revised)
 DT 02-MAR-1990 (first entry)
 XX Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
 DE clones K9-1 through 15e.
 XX Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH.
 XX Hepatitis C virus.
 OS
 FH Key Location/Qualifiers
 FT CDS 3..7310
 FT /*tag= a
 XX
 PN EP318216-A.
 PD
 XX 31-MAY-1989.
 PF 18-NOV-1988; 88EP-00310922.
 XX
 PR 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.
 PR 26-FEB-1988; 88US-00161072.
 PR 06-MAY-1988; 88US-00191263.
 PR 26-OCT-1988; 88US-00263584.
 PR 14-NOV-1988; 88US-00271450.
 XX
 PA (CHIR) CHIRON CORP.
 PA (CHIR) CHIRON CORP.
 XX
 PI Houghton M, Choo QL, Kuo G;
 XX
 DR WPI; 1989-159274/22.
 DR P-PSDB; AAP92050.
 XX
 PT Purified hepatitis C virus - and associated nucleic acids and
 PT polypeptide(s).
 XX
 PS Claim 3; Fig 47-1 - 47-8; 139pp; English.
 XX
 CC It is a double-stranded nucleotide sequence of the open reading frame
 CC (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus
 CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
 CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)
 CC it encodes could be used as immunoassay reagents and vaccines and to
 CC generate antibodies useful in diagnosis and passive immunotherapy for HCV
 CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 1.42e-241 Length: 7310
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 1 Gaps: 0
 US-09-930-591-2 (1-686) x AAN92106 (1-7310)
 Qy 1 MetalloproteaseThra1aTyrAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
 Db 1728 CTGGCGCCCATACGGGGTACGCCACAGACAGAGGGGCTCTAGGGTGCATAATCACC 1787
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValIcLuGlyValGlnIleValSerThra1a 40
 Db 1788 AGCTACTGGCGGGGACAAACCAAGTGGAGGTGAGGTTCAGATTGTGTCNACTGCT 1847
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 1848 GCCCAACCTCTCTGGCAAGCTGCATCATATGGGGTGTGCTGGACTCTCTACCAACGGGGC 1907

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1908 GGAACGAGGACCATCGCGTACCCCAAGGGTCTCTGTATCCAGATGTATACCAATGTAGAC 1967
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1968 CAAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCCTGCATTCG 2027
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 Db 2028 GGCTCTCTGGACCTTACCTGTCACGAGGACGCGGATGTATCTCCGTGGCGCGGG 2087
 QY 121 GlyAspGlyArgGlySerLeuSerProArgProIleSerTyrLeuGlyGlySerSer 140
 Db 2088 GGTGATAGCAGGGGAGCGCTGCTGTGCGCCCGGCCCATTTCTACTTGAAGGGTCTCTCG 2147
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 2148 GGGGGTCCGCTGTGTGCCCCCGGGGACGCGGCGGCATATTTAGGCGCGCGGTGTGC 2207
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 2208 ACCCGTGGAGTGGCTAAGCGGTGACCTTATCCCTGTGGAGAACCTAGACACAACCATG 2267
 QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
 Db 2268 AGGTCCCCGGGTTCACGGATAACTCTCTCCACAGTAGTCCCGCCAGAGCTTCCAGGTG 2327
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 2328 GCTCACTCCATGCTCCACAGGCGGCAAAAGCACCAAGGTCCTCGGTGTGATATGCA 2387
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 2388 GCTCAGGGCTATAAGGTCTAGTACTCAACCCCTCTGTGTGCAACACTGGGGTGTGGT 2447
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 2448 GCTTACATGTCACAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGACAAAT 2507
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 2508 ACCACTGGCAGCCCATCAGTACTCCACCTACGCAAGTTCCTTGGCGAGGGGTGC 2567
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 2568 TCGGGGGCGCTTAAGCATTAATTTGTGACAGTGCCTCCAGGATGCCACATCC 2627
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 2628 ATCTTGGGCATCGGCACCTGCTCTTGACCAAGCAGAGACTGGGGGGGAGACTGTTGTG 2687
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal 340
 Db 2688 CTCGCCACCGCACCCCTCCGGGTCTCGTCACTGTGCCCATCCCAACATCGAGAGGTT 2747
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2748 GCTCTGTCCACCCCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTAATC 2807
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAla 380
 Db 2808 AAGGGGGGAGACATCTCTCTGTCAATTCAAAGAGAGTCCGACGAACTCCCGCA 2867
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 2868 AAGTGTCTGGCATCAATGCGGTGCTTACTACCGCGCTCTTGCCTGCTCGTC 2927
 QY 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2928 ATCCGACCAACGCGATGTTGTCTGTGCAACCGATGCCCTCATGACCGGTATACC 2987
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440


```

Db 2148 GGGGTCGGCTGTTGTCGCCCGGGGACGCGGTGGGCATATTATTAGGCGCGGTGTGC 2207
Qy 161 ThrArgGlyValAlaLysAlaValaspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCCGTGGAGTGGCTAAGCGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 2267
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGGTCCCGGTGTTACGGATAACTCTCTCCACAGTAGTGGCCCGAGAGCTTCCAGGTG 2327
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 2328 GCTCACCTCCATGCTCCACAGGAGCGGCAAAAGACCAAGGTCCCGGCTGCATATGCA 2387
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGGCTTTGGT 2447
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCCAGGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 2507
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCACTGGCAGCCCCATCACGTACTCCACCCTACCGCAAGTTCCTTTCGCGACGGCGGTGC 2567
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGCGCGTATATGACATAATAATTGTGACGAGTGCCACTCCACGGATGCCACATCC 2627
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTGGGCATCGGCACGTGCTCTTGACCAAGCAGACAGACTCGGGGGCGAGACTGGTTGTG 2687
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTGCGCACCGCACCCCTCCGGGCTCGTCACTGTGCCCCATCCACATCGAGGAGTT 2747
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 2807
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCACTCTTGTCATCAAGAGAGAGTGGCAGCACTCGCGCA 2867
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGGTCCGATTGGGCATCAATGCCGTGGCTTACTACCGCGTCTTGACGTGTCGTC 2927
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACCGCGCGATGTTGTCGTGGCAACCGATCCCTTCATGACCGGCTATACC 2987
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCCAGACAGTCGATTTCAGC 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTACCTTCCACATTGACCAATCACGCTCCCGCAGGATGCTGCTCCCGCACT 3107
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 3108 CAACGTCCGGGCGAGCATGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCCGCATGTTCCGACTCGCTCCGCTCTCTGTGAGTGCTATGACGAGGCTGT 3227
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuAlaTyrMetAsnThr 520

```

```

Db 3228 GCTTGCTATAGACTCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 3287
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCCGTGTGCGAGCACCATCTTGAATTTGGAGGGGCTTTTACAGGCCCTC 3347
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCACTTTCTATCCAGACAAAGACAGAGTGGGGAGAACCTTCCCTAC 3407
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 3408 CTGGTAGCGTACCAAGCACCCTGTGCGTAGGGCTCAAGCCCTTCCCCCATCGTGGGAC 3467
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAGTGTTCATTTCGCTCAAGCCACCCTCCATGGGCCAACACACCCTGCTA 3527
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTGGGCGCTGTTTCAGAAATCAATCACCCTGACGACCCAGTCACCAAAATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGACATGCATGTGCGCGCGACTCGAGGTGCTCAGGACACTGGGTGCTGTTGGCGGC 3647
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTCGGCTGCTTTGGCGCGCTATTGCTCTCAACAGGCTGCTGTCATAGTGGCGAG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCTGCTGTCTCGGAGCCGCGCAATCATCTGACAGGAAGTCTCTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 6
AAQ98221
ID AAQ98221 standard; cDNA to mRNA; 7310 BP.
AC AAQ98221;
XX
XX 25-MAR-2003 (revised)
DT 15-AUG-1996 (first entry)
XX
XX Hepatitis C virus clone genome.
XX Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;
XX antibodies; immunoprophylaxis; sera; serum; ds.
XX Hepatitis C virus.
OS
XX US5443965-A.
FN
XX 22-AUG-1995.
PD
XX 05-APR-1991; 91US-00681703.
PF
XX 06-APR-1990; 90US-00505611.
PR 09-OCT-1990; 90US-00594854.
XX (GENE-) GENELABS INC.
PA
XX Moeckli R, Reyes GR, Kim JP;
XX
XX WPI; 1995-302120/39.
DR
XX New nucleic acids encoding hepatitis C virus antigens - used to develop
PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti
PT -HCV antibodies.
XX

```

Example 4: Fig 11: 71pp: English:

Hepatitis C virus (HCV) antigens can be used for detecting HCV infected sera and individuals infected with HCV. They can also be used in an anti-HCV vaccine or for the production of anti-HCV antibodies which can be used for passive immunoprophylaxis. The antigens consistently identify more HCV positive serum samples with a high degree of specificity. See CC AAR96203-14 and AAR81939-51. (Updated on 25-MAR-2003, to correct PF field.) (Updated on 25-MAR-2003 to correct PR field.)

Sequence 7310 BP: 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.42e-241	Length:	7310
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	2	Gaps:	0

US-09-930-591-2 (1-686) x AAQ98221 (1-7310)

Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	1728	CTGGCGCCCATACGGGGTACGCCACAGACAAGGGGCTCTCCPAGGGTGCAATAACACC	1781
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla	40
Db	1788	AGCCTAACTGGCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGCAACTGCT	1847
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
Db	1848	GCCCAAAACCTTCTCTGGCAAGTGATCAATGGGGTGCTGGACTGTCTACCAACGGGGCC	1907
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	1908	GGACACGAGGACCATCGGGTCACCCAGAGGGCTCTGTATCCAGATGTATACCAATGTAGAC	1967
Qy	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	1968	CAAGACCTTGTGGCTGCGCCGCTCCGCAAGTAGTACCGCTCATTTGACACACCTGTGACATTGC	2027
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
Db	2028	GGCTCTCTCGGACCTTTACCTTGTCACGAGGACACGCCGATGTCTCCGTCGGCGCGGG	2087
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	2088	GGTGATAGCAGGGCAGCGCTCTGTGCCCCCGGCCCATTTTCTACTTTGAAGAGGCTCTCG	2147
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	2148	GGGGGTTCGCTTGTTGTGCCCCCGCGGGCACGCCGTGGGCATATTTAGGCGCGCGGTGTC	2207
Qy	161	ThrArgGlyWalAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	2208	ACCCGTGGAGTGGGTAAAGGCGGTGGACTTTATCCGTGTGGAGAACCCTAGACAAACCATG	2267
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	2268	AGGTCCCGGGTGTTCCCGGATTAATCTCTCTCCACCCAGTAGTGTGCCCCAGAGTTCACGGTG	2327
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	2328	GCTCACCTCCATGTCTCCACAGGACGCGGCAAAAGACCAAGGTCCCGGCTGCATATGCA	2387
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	2388	GCTCAGGGCTATAAGTGCTTAGTACTCAACCCCTCTGTGCTGCACACTCGGGCTTTGGT	2447
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	2448	GCTTACATGTCGAAGGCTCATGGATCGATCCTAACATCAGACCGGGGTGAGAACCAATT	2507

QY	261	ThrThrGlySerProIleThrThrTySerThrTyGlyIysPheLeuAlaAspGlyGlyCys	280
DB	2508	ACCACCTGGCAGCCCCATCAGTACTCCACTACGGCAAGTTCCTTCGCCACGCGGGTGCC	2567
QY	281	SerGlyGlyAlaTyzAspIlelleleleleCysAspGluCysHisSerThrAspAlaThrSer	300
DB	2568	TCGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCACCTCCACGGATGCCACATCC	2627
QY	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
DB	2628	ATCTTGGGCGATCGGCACTGTCTTGACCAAGCAGACACTCGCGGGCGGAGACTGGTTGTG	2687
QY	321	LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal	340
DB	2688	CTCGCCACCGCCACCCTTCGGGCGCTCCGTCACTGTGCCCGCATCCCAACATCGAGGAGGTT	2747
QY	341	AlaLeuSerThrThrGlyGluIleProPheTyGlyIysAlaIleProLeuGluAlaIle	360
DB	2748	GCCTCTGCCACCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC	2807
QY	361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
DB	2808	AAAGGGGGGAGACATCTCATCTTCTGTCATTCAAAGAAGATGCGACGAATCGCGCCA	2867
QY	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyzTyrArgGlyLeuAspValSerVal	400
DB	2868	AACTGCTGCATGGGCATCAATGCGTGGCTACTACCGCGGCTCTTGACGTGTCCGTC	2927
QY	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
DB	2928	ATCCGACACGCGCGCATGTTGTCGTCTGGCAACCGATGCCCTCATGACGCGCTATACC	2987
QY	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
DB	2988	GGCGACTTCGACTCGGTGATAGACTGCATACGTGTGTCAACCAGACAGTCTGATTTCAGC	3047
QY	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
DB	3048	CTTGACCTACCTTACCATTGAGACAATCACGCTCCCCAGAGATGCTGTCTCCGCGACT	3107
QY	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
DB	3108	CAACGTCGGGGCAGGACTGGCAGGGGAGCCAGGCACTTACAGATTGTGGCACCCGGGG	3167
QY	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
DB	3168	GAGCGGCCCTCCGCGCATGTCGACTGTCGCTCTCTGTGAGTGCTATGACGAGGCTGT	3227
QY	501	AlaTrpTyzGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
DB	3228	GCCTTGATAGCTCACCCCGCCGAGACTACAGTTAGGTACTCGAGCGTACATGAACACC	3287
QY	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
DB	3288	CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCGTC	3347
QY	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
DB	3348	ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCCTTAC	3407
QY	561	LeuValAlaTyzGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp	580
DB	3408	CTGCTAGCGTACCAAGCCACCGTGTGGCTTAGGGCTCAAGCCCTCCCTCCATCGTGGGAC	3467
QY	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
DB	3468	CAGATGTGGAGTGTGTGATTCGCTCTCAAGCCACCCCTCCATGGGCCAACCCCTGCTA	3527
QY	601	TyzArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyzIle	620
DB	3528	TACAGACTGGGCGGTGTCAAGATGAATGAATCACTGTACCGACCCAGTCAACAAATACATC	3587

QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 DB 3588 ATGACATGCTGGCGCGACCTGGAGGTCTGACGAGCACCCTGGTGTCTGTTGGCGGC 3647
 QY 641 ValLeuAlaLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 660
 DB 3648 GTCTGTGGTGTCTGGCGCGCTATTTGCTGTCAACAGGCTGCGTGTGTCATAGTGGCGAGG 3707
 QY 661 IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuValArgGluPhe 680
 DB 3708 GTGCTGTCTGGAGACCGCAATATACCTGACAGGGAAGTCTCTACCGAGAGTTC 3767
 QY 681 AspGluMetGluGluCys 686
 DB 3768 GATGAGATGGAAGAGTGC 3785
 RESULT 7
 AAA75296
 ID AAA75296 standard; cDNA; 8316 BP.
 AC AAA75296;
 DT 15-JAN-2001 (first entry)
 DE cDNA sequence compiled Hepatitis C virus cDNA clones.
 KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
 KW viral infectivity; viral replication; ds.
 OS Hepatitis C virus.
 FH Key Location/Qualifiers
 FT CDS 1..8316
 FT /*tag= a
 FT /note= "partial sequence; no termination codon given"
 XX
 XX EP1034785-A2.
 XX 13-SEP-2000.
 XX 16-MAR-1990; 2000EP-00109602.
 XX 17-MAR-1989; 89US-00325338.
 XX 20-APR-1989; 89US-00341334.
 XX 18-MAY-1989; 89US-00355002.
 XX 16-MAR-1990; 90EP-00302866.
 XX (CHIR) CHIRON CORP.
 XX Houghton M, Choo Q, Kuo G;
 XX WPI; 2000-566891/53.
 XX P-FSDB; AAB18540.

Novel composition comprising a hepatitis C virus antisense polynucleotide which is complementary to or corresponds to a sense strand of the virus genome, and selectively hybridizes to it.

Example; Fig 16; 75pp; English.

The specification describes a pharmaceutical composition which comprises a hepatitis C virus (HCV) antisense polynucleotide. The HCV is characterized by a positive stranded RNA genome which has 40% homology at the polypeptide level to a HCV polyprotein. The antisense polynucleotide binds to cellular polynucleotides which enhance and/or are required for viral infectivity, replicative ability or chronicity. The antisense polynucleotides may also be designed to bind with high specificity, to be of increased stability, to be stable and to have low toxicity. The composition also comprises an agent which causes viral RNA to be inactive. The composition is used for preventing HCV replication in a system. The present sequence represents a novel HCV cDNA sequence, which is used in the course of the invention

SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.66e-241 Length: 8316
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 3 Gaps: 0
 US-09-930-591-2 (1-686) x AAA75296 (1-8316)
 QY 1 MetAlaProIleThrAlaIleThrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 2734 CTGGCGCCCATCAGCGCGTAGCCCGCAGCAGCAAGGGCCCTCTAGGTGTCATATCACC 2793
 QY 21 SerLeuThrGlyArgAspIleAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 2794 AGCCTAACTGGCCGGGCAAAAACCAAGTGAGGTGAGGTCCAGATGTGTCAACTGCT 2853
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThrHisGlyAla 60
 DB 2854 GCCCAACCTTCTGGCAACGTCATCAATGGGGTGTGCTGACTGCTACCCGGGGCC 2913
 QY 61 GlyThrArgThrIleAlaSerProIleGlyProValIleGlnMetThrThrAsnValAsp 80
 DB 2914 GGAACGAGGACCATCGCGTCACCAAGGGTCTCTGTCATCCAGATGTATACCAATGTAGAC 2973
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 2974 CHAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCGCTCATTTGACACCTGCTGCTGC 3033
 QY 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArg 120
 DB 3034 GGCCTCTCGGACCTTTACTGCTCAGGAGCAGCCGATGTCTATCCCGTGGCGCGCGG 3093
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerThrLeuLysGlySerSer 140
 DB 3094 GGTGATAGCAGGGCGAGCTGCTGCGCCCGCGCCCATTTCTTACTTGAAGGCTTCTCTCG 3153
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 3154 GGGGGTCCGCTGTTGTGCGCGCGGCGCAGCCCTGGGCATATTTAGGGCGCGGTGTGC 3213
 QY 161 ThrArgGlyValAlaAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
 DB 3214 ACCCGTGGAGTGGCTAAGCGGTGACTTTATCCTGTGGAGAACCTAGACACCAACATG 3273
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerThrGlnVal 200
 DB 3274 AGTCCCCCGGTGTTACGGATAACTCTCTCCACAGTAGTGGCCCGCAGAGCTTCCAGGTG 3333
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaThrAla 220
 DB 3334 GCTCACTTCCATGCTCCACAGGCGGCAAAAGCAGGTCCTCCGCGTGCATATGCA 3393
 QY 221 AlaGlnGlyThrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 3394 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGTCGAACACTGGGCTTGGT 3453
 QY 241 AlaThrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 3454 GCTTACATGTCCAAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACATTT 3513
 QY 261 ThrThrGlySerProIleThrThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 DB 3514 ACCACTGGCAGCCCATCAGTACTTCCACTACCGCAAGTCTCTTGGCGAGCGCGGTGC 3573
 QY 281 SerGlyGlyAlaThrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 3574 TCGGGGGCGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGGATGCCACATCC 3633
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320

Db 3634 ATCTTGGGCAATCGGCACTGCTTGGCAAGAGAGACTCGGGGGGAGACTGGTGTG 3693
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 3694 CTGCCACCCGCCACCTCCGGCTCGCTCACTGTGCCCATCCCAACATCGAGGAGGTT 3753
Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyGlyAlaIleProLeuGluAlaIle 360
Db 3754 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 3813
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 3814 AAGGGGGGAGACATCT 3873
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaThrTyrTyrArgGlyLeuAspValSerVal 400
Db 3874 AAGTGTGTGTCATTTGGGCATCAATGCGGTGCTACTACCGCGTCTTGACGTGTCGTC 3933
Qy 401 IleProThrSerGlyAspValValValValAlaThrThrAspAlaLeuMetThrGlyPheThr 420
Db 3934 ATCCCGACAGCGGCGATGTTGCTGCTGGCAACCGATGCCCTCATGCCGCTATACC 3993
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 3994 GGGCACTTCGACTCGGTGATAGACTGCATATACGTGTGTCAACCGACAGTCGATTCAGC 4053
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4054 CTGACCTCTACCTTCACCATTTAGACAACTACGCTCCCCAGAGTGTCTCCGCACT 4113
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4114 CAACGTCGGGCGAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGACCGGG 4173
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4174 GAGCGCCCTCCGCACTGTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4233
Qy 501 AlaThrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4234 GCTTGGTATGACTCAGCCCGCGAGACTACAGTATAGGTACGAGGATACATGAACACC 4293
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 4294 CCGGGGCTCCGCTGTCCAGGACATCTTGAATTTGGGAGGGCGTCTTTACAGGCGTC 4353
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 4354 ACTCATATAGATGCCACTTCTATCCAGAACAAAGCAGAGTGGGGAGAACCTTCTTAC 4413
Qy 561 LeuValAlaThrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 4414 CTGGTAGCGTACCAAGCACCGCTGTGGCTAGGGCTCAAGGCCCTCCCCCATCGTGGAC 4473
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 4474 CAGATGTGAAGTGTGTGATTCGGCTCAAGCCCACTCCATGGGCGCCCAACCCCTGCTA 4533
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 4534 TACAGACTGGGCGCTGTTTCAAGTGAATCACTCAGCCAGCCAGCCAGCCAGCCAGCC 4593
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640
Db 4594 ATGACATGATGTGGCGGACCTGGAGTGTCTACGAGCACCTGGGTGCTCGTGGCGGC 4653
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 4654 GTCTGGCTGTGTGGCGCGGATTTGCTGTACACAGGCTGCGTGTGTGTGTGTGTGT 4713
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
:::

Db 4714 GTCGTCTTGTCCGGGAAGCCGGCAATCATCTACAGGGAAGTCTCTACGAGAGTTC 4773
Qy 681 AspGluMetGluGluCys 686
Db 4774 CATGAGATGGAAGAGTGC 4791
RESULT 8
AAZ07656
ID AAZ07656 standard; DNA; 9133 BP.
XX AAZ07656;
XX 20-MAR-2003 (revised)
DT 08-NOV-1999 (first entry)
XX Nucleotide sequence of HCV-1 ORF.
XX Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;
KW HCV infection; vaccine; ds.
XX Hepatitis C virus.
PH Key Location/Qualifiers
FT CDS 268..9132
FT /tag= b
FT /transl_except= (pos:1588..1589; aa:Leu)
FT /note= "this codon has an apparent 1 nucleotide deletion,
FT which alters the reading frame"
FT /transl_except= (pos:1647..1650; aa:Pro)
FT /note= "this codon has an apparent 1 nucleotide
FT insertion, which alters the reading frame; this insertion
FT is not indicated in the sequence present in the formal
FT sequence listing of the specification"
XX EP939128-A2.
XX 01-SEP-1999.
XX 17-SEP-1990; 99EP-00101746.
XX 15-SEP-1989; 89US-00408045.
XX 21-DEC-1989; 89US-00456142.
XX 17-SEP-1990; 90EP-00310149.
XX (OYAA/) OYA A.
PA (CHIR) CHIRON CORP.
XX Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;
PI Cha T, Irvine BD;
XX WPI; 1999-480843/41.
DR P-PSDB; AAY14975.
XX New Hepatitis C Virus isolates, useful for diagnosis of hepatitis
PT infections and development of vaccines.
XX Disclosure; Fig 12; 132pp; English.
XX The invention provides two new isolates of hepatitis C virus (HCV), J1
CC and J7. These two isolates comprise nucleotide and amino acid sequences
CC that are distinct from the HCV isolate HCV-1. The nucleotide sequences
CC may be used to detect non-A, non-B HCV (NANBH) polynucleotides by
CC hybridization for diagnosis of NANBH infections. They may also be used to
CC screen blood donors, donated blood and blood products for this infection.
CC The isolates may also be used to isolate other naturally occurring
CC variants of the virus. The polypeptides may be used as a vaccine for
CC administration to patients to protect against infection with NANBH. The
CC present sequence represents the nucleotide sequence of HCV-1 ORF.
CC (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC correct PR field.)
XX Sequence 9133 BP; 1834 A; 2772 C; 2600 G; 1927 T; 0 U; 0 Other;
SQ

Alignment Scores:		Pred. No.:	1,866-241	Length:	9133
		Score:	3574.00	Matches:	672
		Percent Similarity:	99.42%	Conservative:	10
		Best local Similarity:	97.96%	Mismatches:	4
		Query Match:	98.78%	Indels:	0
		DB:	2	Gaps:	0
US-09-930-591-2 (1-686) x AA207656 (1-9133)					
QY	1	MetAlaProLeuThrAlaThrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20		
DB	3343	CTGGCGCCCATCACGGGTACGCCAGCAGACAAAGGGGCTCTTAGGGTGCAATAATCACC	3402		
QY	21	SerLeuThrGlyArgAspLeuValGluGlyGluValGlnIleValSerThrAla	40		
DB	3403	AGCCTAACTGCGCGGACAAACCAAGTGAGGGTGAGGTCCAGATTGTCACTGCT	3462		
QY	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60		
DB	3463	GCCCAAACTTCTTGGCAACGTGCATCAATGGGGTGTGCTGCACTGTCTACCAAGGGGCT	3522		
QY	61	GlyThrArgThrIleAlaSerProGlySerProValIleGlnMetTyrThrAsnValAsp	80		
DB	3523	GGAAGGAGACCATCGCGTCACCAAGGGTCTCTGCATCCAGATGTATACCAATGTAGAC	3582		
QY	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100		
DB	3583	CAAGACCTTGTGGCTGGCGGCTCCGCAAGGTAGCGGCTCATTTGACACCTGCACCTTCG	3642		
QY	101	GlySerSerAspLeuThrValThrArgHisAlaAspValIleProValArgArgArg	120		
DB	3643	GGCTCCTCGGACCTTACCTGGTCACGAGGACCGCCGATGTCTCCGTCGCGCGGCGG	3702		
QY	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer	140		
DB	3703	GGTGATGACAGGGGACGCTGCTGTCGCCCGGCCCATTTCTACTTGAAGGCTCCTCG	3762		
QY	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160		
DB	3763	GGGGGTCCGCTGTGTGCCCGCGGGGACGCGCGTGGGCATATTTAGGGCGCGGGTGTGC	3822		
QY	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180		
DB	3823	ACCCGTGGAGTGGCTAAGCGGTGAGATTATCTCTGTGGAGAACCTAGAGACCAATG	3882		
QY	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200		
DB	3883	AGGTCCCGGGTGTTCAGCGATACTCTCTCCACAGTAGTGGCCCAAGAGCTTCCAGGTG	3942		
QY	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220		
DB	3943	GCTCACCTCCATGTCTCCACAGGCGGCGGCAAGACCAAGGTCCCGGTGCTGATGCA	4002		
QY	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240		
DB	4003	GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGCTGCTGCAACACTGGGCTTTGGT	4062		
QY	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260		
DB	4063	GCTTACATGTCCAAAGGTCTAGGATCGATCTCTTAACATCAGGACCGGGGTGAGAACAT	4122		
QY	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280		
DB	4123	ACCACTGGAGCCCATCACGTACTCCACTACGGCAAGTTCCTTGGCGAGCGGGGTGC	4182		
QY	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300		
DB	4183	TCGGGGGGCGGTATATGACATAAATAATTGTGAGAGTGCCACTCCACGGATGCCACATCC	4242		
QY	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal	320		
DB	4243	ATCTTGGGCATCGGCATGTCTCTTGACCAAGCAGAGACTCGCGGGGGAGACTGTTGTG	4302		
QY	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal	340		
DB	4303	CTCGCCACCGCCACCCCTCCGGCTCCGTCCTGCTGTCGCTCCCAACATCGAGAGGTT	4362		
QY	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360		
DB	4363	GCTCTGTCCACCCAGGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAGTAATC	4422		
QY	361	LysGlyGlyArgHisIlePheCysHisSerLysLysCysAspGluLeuAlaIle	380		
DB	4423	AAGGGGGGAGACATCTCTCTTCTGTCATTAAGAAAGAGTGGAGAACTCCGCCCA	4482		
QY	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400		
DB	4483	AAGCTGGTGGCATTTGGGCATCAATGCGTGGCTACTACCGGGGTCTTGACGTGCTCGTC	4542		
QY	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420		
DB	4543	ATCCCGACCGAGCGGATGTTGTCGTCGTGGCAACCGATGCCCTCATGACCGGTATACC	4602		
QY	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440		
DB	4603	GGCGACTTCGACTCGGTGATAGCTGCAATAGTGTGTACCCAGACAGTGCATTTGACG	4662		
QY	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460		
DB	4663	CTTGACCTTACCTTCCACATTTGAGACAATCACGCTCCCGCAGGATGCTGTCTCCGCACT	4722		
QY	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480		
DB	4723	CAACGTCCGGGAGGAGCTGGCAGGGGAGCCAGCATCTACAGATTGTGGCACCCGGG	4782		
QY	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500		
DB	4783	GAGCGCCCTCCGGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4842		
QY	501	AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520		
DB	4843	GCTTGGTATAGCTCACCGCCCGGAGACTACAGTTAGGCTTACGAGCGGTATGAAACAC	4902		
QY	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540		
DB	4903	CCGGGGCTTCCGTTGTCAGACCATCTTGAATTTGGGAGGGGCTTTTACAGGCTC	4962		
QY	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560		
DB	4963	ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC	5022		
QY	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp	580		
DB	5023	CTGTGTAGGTATACCAAGCCCGCTGTCGCTAGGCTCAAGCCCTCCCTCCCATCTGGGAC	5082		
QY	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600		
DB	5083	CAGATGTGAAAGTGTGTTGATTCGCTCAAGCCCAACCTCCATCGGGGCAACACCCCTGTA	5142		
QY	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620		
DB	5143	TACAGCTGGGGCTGTGTTGAGATGAATACCTCTGAGCCACCCAGTCCCAATATCATC	5202		
QY	621	MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly	640		
DB	5203	ATGACATCATGTCCGGCGACCTGGAGTCTGTCAGACACCTGGGTGCTGTTGGCGGC	5262		
QY	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValIleValGlyArg	660		
DB	5263	GTCTGTGGCTGTTTGGCGCGGTATTTGCTGTCAACAGCTGCGGTGTCATAGTGGGAGG	5322		
QY	661	IleValLeuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe	680		
DB	5323	GTCTGTGTTTGGCGGAGCGGCAATCATCTGACAGGGAAGTCTCTTACCGAGAGTTC	5382		


```

QY 681 AspGluMetGluGluCys 686
Db 5383 CATGACATGGAAGAGTGC 5400

RESULT 9
AAQ10566
ID AAQ10566 standard; DNA; 9185 BP.
XX AC AAQ10566;
XX 25-MAR-2003 (revised)
DT 29-APR-1991 (first entry)
XX
DE Hepatitis C virus strain 1 DNA.
XX
KW Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;
KW viral infections; ss.
XX
OS Hepatitis C virus.
XX
PN EP414475-A.
XX
PD 27-FEB-1991.
XX
PF 21-AUG-1990; 90EP-00309120.
XX
PR 25-AUG-1989; 89US-00398667.
XX (CHIR ) CHIRON CORP.
XX
PA Weiner AJ, Steimer KS;
PI WPI; 1991-059670/09.
XX
DR Cell lines infected with hepatitis C virus - are used as source of
PT antigens for detection of HCV antibodies, for vaccines, and for screening
PT anti-viral agents.
XX
PS Disclosure; Fig 1; 24pp; English.
XX
CC This is a hepatitis C virus (HCV) composite cDNA sequence, deduced using
CC overlapping clones. a compsn. contg. the antigenic protein encoded by
CC this sequence is useful for detecting anti-HCV anti- bodies (Abs) and for
CC screening an agent which inhibits HCV replic- action. A cell line infected
CC with this virus can be used as a source of antigens. The antigen is
CC useful for preparing vaccines for treating viral infections. See also
CC AAQ10567. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.88e-241 Length: 9185
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 2 Gaps: 0

US-09-930-591-2 (1-686) x AAQ10566 (1-9185)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3395 CTGGCGCCCATCCAGCGGTAGCCGACAGACAAGGGCCCTCTAGGGTGCATAATCACG 3454
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3455 AGCCTAACTGGCCGGGCAAAAACCAAGTAGTGGAGGTGAGTCCAGATTGTGCAACTGCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysIleThrValTyrHisGlyAla 60
Db 3515 GCCCAACCTTCTCGGCAACGTGTCATCAATGGGGTGTGCTGGACTGTCTACCAAGGGGCC 3574
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

```

Db 4655 GCGACITTCGACTCGGTGATAGACTGCAATACGTGTGTCCACCCAGACAGTCGATTTCAGC 4714
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4715 CTTGACCTTACCTTCACCAATTGACCAATCACGCTCCCCAGGATGCTGTCTCCCGCACT 4774
 Qy 461 GlnArgArgGlyArgThrGlyValGlyArgProGlyIleThrValArgPheValAlaProGly 480
 Db 4775 CAAGCTCGGGCAGACTGCGCAGGGGAGAGCCAGGCACTACAGATTGTGGACACCGGG 4834
 Qy 481 GluArgProSerGlyMetPheAspSerSerValuGluCysGluCysThrAspAlaGlyCys 500
 Db 4835 GAGCGCCCTCCGCGCATGTTCCGACTCGCTCGCTCTGTGAGTGCTATGACGCGAGGTGT 4894
 Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 4995 GCTTGGTATGAGCTCAGCCCGCCGAGACTACAGTTAGGCTACGAGGCTACATGAACACC 4954
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 4955 CCGGGCTTCCCGTGTGCGAGGACCACTTGAATTTTGGAGGGCGTCTTTACAGGCTC 5014
 Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrIleThrIleSerGlyGluAsnLeuProTyr 560
 Db 5015 ACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTAC 5074
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 5075 CTGCTAGCTACCAAGCAGCGTGTGGCTAGGGCTCAAGCCCTCCCATCGTGGGAC 5134
 Qy 581 GlnVetTrpIysCysLeuIleArgLeuIysProThrLeuHisGlyProThrProLeuLeu 600
 Db 5135 CAGATGTGGAGTGTGTTGATTCGCTCAAGCCCAAGCCCTCCATGGGCAACACCCCTGCTA 5194
 Qy 601 TyrArgLeuGlyAlaValAlaGlnAspGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 5195 TACAGATGGCGCTGTTCAGATGAATCACCTTGACGACCCAGTCACCAATACATC 5254
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 5255 ATGACATGCTATCGGCGGACCTGGAGTGGTCAAGGACCTGGGTGCTCGTGGCGG 5314
 Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrCysValValIleValGlyArg 660
 Db 5315 GTCTGCTGCTTTGGCGCGGTATGCTGTCAACAGGCTGCGTGTGTCATAGTGGCAGG 5374
 Qy 661 IleValLeuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 5375 GTCTGCTTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGATTTC 5434
 Qy 681 AspGluMetGluGluCys 686
 Db 5435 GATGAGATGAAGAGTGC 5452

RESULT 10

AAA75297

ID AAA75297

AC AAA75297;

XX AAA75297;

DT 15-JAN-2001 (first entry)

XX Sense strand of HCV encoding a polyprotein.

XX Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;

XX viral infectivity; viral replication; ds.

XX Hepatitis C virus.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX EP1034785-A2.
 XX 13-SEP-2000.
 XX 16-MAR-1990; 2000EP-00109602.
 XX 17-MAR-1989; 89US-00325338.
 XX 20-APR-1989; 89US-00341334.
 XX 18-MAY-1989; 89US-00355002.
 XX 16-MAR-1990; 90EP-00302866.
 XX (CHIR) CHIRON CORP.
 XX Houghton M, Choo Q, Kuo G;
 XX WPI; 2000-566891/53.
 XX P-PSDB; AAB18541.
 XX Novel composition comprising a hepatitis C virus antisense polynucleotide
 PT which is complementary to or corresponds to a sense strand of the virus
 PT genome, and selectively hybridizes to it.
 XX Example; Fig 17; 75pp; English.
 XX The specification describes a pharmaceutical composition which comprises
 CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
 CC characterized by a positive stranded RNA genome which has 40% homology at
 CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide
 CC binds to cellular polynucleotides which enhance and/or are required for
 CC viral infectivity, replicative ability or chronicity. The antisense
 CC polynucleotides may also be designed to bind with high specificity, to be
 CC of increased stability, to be stable and to have low toxicity. The
 CC composition also comprises an agent which causes viral RNA to be
 CC inactive. The composition is used for preventing HCV replication in a
 CC system. The present sequence represents a novel HCV cDNA sequence, which
 CC is used in the course of the invention

SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.88e-241	Length:	9185
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	3	Gaps:	0

US-09-930-591-2 (1-686) x AAA75297 (1-9185)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 3395 CTGGCGCCCATCACGGCGGTACGCCAGACAGAGGGGCGCTCTAGGGTGATAATCACC 3454
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 3455 AGCCTACTGTCGGCGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTCACTGCT 3514
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 3515 GCCCAACCTTCTCGGCAACGTGCATCAATGGGCTGTCTGAGCTGTCTACCGGGGCC 3574
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 3575 GGAACGAGGACCATCGGTACCCCAAGGTCCTGTATCCAGATGTATACCAATGTAGAC 3634
 Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 3635 CAAGACCTTGTGGGCTGCGCCGCTCCCAAGGTAGCGGCTCATTGACACCCCTGCACCTGC 3694
 Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 Db 3695 GCCTCTCGGACCTTTTACCTGGTCCAGGACGCCGATGTCTATCCCGTCCCGCGCGG 3754

Key Location/Qualifiers
 CDS 320..9184
 /*tag= a
 /note= "partial sequence; no termination codon given"

```

QY 121 GluAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 3755 GGTGATAGCAGGCGACCTGCTGTCGCCCGCGGCCATTTCTTACATTGAAGGCTCTCTCG 3814
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3815 GGGGGTCCGGCTGTGTGTCGCCCGCGGGCACCGCGTGGGCATATTTAGGGCCGCGGTGTC 3874
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3875 ACCCGTGGAGGGGTAGAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3935 AGTCCCGCGGTTCACGGATAACTCTCTCCACCACTAGTGGCCCGACAGCTTCCAGGTG 3994
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3995 GCTCACCTCCATGCTCCACAGGCGAGCGGAAAGCACCAGGTCGCCGCTGCATATGCA 4054
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4055 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCACACACTGGGCTTTGGT 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaArgThrIle 260
DB 4115 GCTTACATGTCCAAAGGCTATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAAT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 4175 ACCACTGGCAGCCCATCAGTACTCCACTACGGCAAGTTCCTTGGCCGAGCGGGGTGC 4234
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 4235 TCGGGGGCGGCTATGACATAATAATTTGTGACGAGTGCACCTCCACGGATGCCACATCC 4294
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4295 ATCTTGGGCATCGGCATGTCTTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTGTG 4354
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4355 CTGCGCACCGCCACCTCCGGGCTCGTCACTGTGCCCCATCCCAACATCGAGGAGGTT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 4415 GCTCTGTCCACCCAGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 4474
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 4475 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAGAGAAGTGGCGAGACTCGCGCA 4534
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4535 AAGTGTGTCGATTTGGCATCAATCGGTGGCTACTTACCGCGTCTTTGACGTCCGTC 4594
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 4595 ATCCCGACCAAGCGGAGATGTGCTGTGCGAACCATGCTCCCTCATGACCGGCTATACC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4655 GGCACATTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCGATTTCCAGC 4714
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4715 CTTGACCGCTACCTTCCACATTGAGACAATCAGCTCCCGCAGGATGCTGTCTCCCGCACT 4774
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 4775 CAACGTGGGGCAGGATGTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG 4834

```

```

QY 481 GluAspProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 4835 GAGCGCCCTCCGGCATGTTTCGACTCTCGTCCGCTCTCTGTGAGTGTATGACGCGGCTGT 4894
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4895 GCTTGGTATGAGTCAACGCCCGCCGAGACTACAGTTAGGCTACGAGCGGTACATGAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
DB 4955 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGGCTTTTACAGGCTTC 5014
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 5015 ACTCATATAGATGCCACTTCTTATCCACAGCAAAAGCAGAGTGGGAGAACCTTCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB 5075 CTGGTAGGCTACCAAGCACCGGTGTGCGTAGGGCTCAAGCCCCCTCCCCCATGTGGGAC 5134
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGTGGAAGTGTTCGATTCGCTCAAGCCCCACCTCCATGGGCCCAACACCTTCTTAC 5194
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 5195 TACAGACTGGGCGCTGTTTCAAGTAAGAAATCAACCTGCACGCCACCCAGTACCAAAATACATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
DB 5255 ATGACATCAATGTCCGCGGACCTGGAGGTGCTCAGCAGACACTGGGTGCTGTTGGCGGC 5314
QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 5315 GTCTCGGCTGCTTTGGCGCGCTATTTCCTGTCAACAGGCTGCGTGTCTATAGTGGCAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 5375 GTCTGCTTGTCCGGAAGCCGCAATCATCTGACGGGAAGTCTCTTACCGAGAGTTC 5434
QY 681 AspGluMetGluGluCys 686
DB 5435 GATGAGATGGAAGAGTGC 5452

RESULT 11
AAT12710
ID 'AAT12710 standard; cDNA; 9401 BP.
XX AC AAT12710;
XX XX
DT 25-MAR-2003 (revised)
DT 15-MAY-1996 (first entry)
XX XX
DE Hepatitis C virus polyprotein.
XX XX
KW Non-A non-B hepatitis virus; NANBH; HCV; antigen; detection; diagnosis;
KW antibodies; ds.
XX OS Hepatitis C virus.
XX FH Key Location/Qualifiers
FT CDS 342..9378
FT /tag= a
XX PN EP693687-A1.
XX XX
PD 24-JAN-1996.
XX PF 03-APR-1991; 95EP-00114016.
XX XX
PR 04-APR-1990; 90US-00504352.
XX XX
PA (CHIR ) CHIRON CORP.

```



```

Db      5097 CTGGTAGGTTACCAAGCCACCGTGGCTAGGCTCAAGCCCTCCCCCATCGTGGAC 5156
QY      581 GlnMetThrLysCysLeuLeuAArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      5157 CAGATGTGGAAGTGTTCATTTCGCTCAAGCCACCCCTCCATGGCGCCACACCCCTGCTA 5216
QY      601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      5217 TACAGCTGGGGCTGTTCAGAAATGAATCAACCTGAGCCAGCCAGTCACCAATACATC 5276
QY      621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db      5277 ATGACATCATGTCGGCCGACCTGGAGTCTGTCACGACCTGGTGTCTGTGGCGGC 5336
QY      641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      5337 GTCTGTGGCTGTTCGGCGGTATTGCTGTCAACAGCTGCGTGGTCATAGTGGCAGG 5396
QY      661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db      5397 GTCTGTGTTCGGGAGCCGCAATCATCTGACAGGGAGTCTCTACCGAGATTC 5456
QY      681 AspGluMetGluGluCys 686
Db      5457 GATGAGATGGAAGAGTGC 5474

RESULT 12
AAT99981
ID      AAT99981 standard; DNA; 9401 BP.
XX      AAT99981;
AC      XX
AC      XX
DT      25-MAR-2003 (revised)
DT      16-MAR-1998 (first entry)
DE      HCV polypeptide coding sequence.
XX      PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW      C domain; S domain; NS; HCV polypeptide; anti-HCV antibody; detection;
KW      NS4; ds.
XX      Hepatitis C virus.
XX      Key Location/Qualifiers
FT      CDS 342..9377
FT      /*tag= a
XX      US5683864-A.
XX      04-NOV-1997.
XX      07-JUL-1992; 92US-00910760.
XX      18-NOV-1987; 87US-00122714.
XX      30-DEC-1987; 87US-00139886.
XX      26-FEB-1988; 88US-00161072.
XX      06-MAY-1988; 88US-00191263.
XX      26-OCT-1988; 88US-00263584.
XX      14-NOV-1988; 88US-00271450.
XX      17-MAR-1989; 89US-00325338.
XX      20-APR-1989; 89US-00341334.
XX      21-APR-1989; 89US-00353896.
XX      18-MAY-1989; 89US-00355002.
XX      04-APR-1990; 90US-00504352.
XX      (CHIR ) CHIRON CORP.
XX      Kuc G, Houghton M, Choo Q;
XX      WPI; 1997-548976/50.
XX      P-PSDB; AAW34480.
XX

```

```

PT      Combination of three hepatitis C virus antigens - used for detection of
XX      specific antibodies to diagnose infection.
XX      Disclosure; Col 25-46; 57pp; English.
XX
CC      This sequence represents the Hepatitis C virus polypeptide coding
CC      sequence. Fragments of this sequence can be amplified and used in the
CC      combination of HCV antigens of the invention. The HCV antigen combination
CC      comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
CC      1-120 of the HCV polypeptide), or its immunologically reactive fragment
CC      containing at least 8 aa. It also comprises two additional antigens from
CC      two different polypeptide domains, including at least 8 aa from the NS3,
CC      NS4, S or NS5 domains of the polypeptide, corresponding, respectively, to
CC      aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polypeptide.
CC      Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
CC      regions of the HCV polypeptide. These antigen combinations are used
CC      diagnostically to detect anti-HCV antibodies, using any standard
CC      immunoassay format. These antigen combinations have a broader range of
CC      reactivity with antibodies than any antigen individually. (Updated on 25-
CC      MAR-2003 to correct PR field.)
XX
SQ      Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

```

Alignment Scores:

```

Pred. No.: 1 93e-241 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 2 Gaps: 0

```

US-09-930-591-2 (1-686) x AAT99981 (1-9401)

```

QY      1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleLeuThr 20
Db      3417 CTGGCGCCCATCACGGCGTACGCCAGCAGCAGAGGGGCTCTCTAGGTGCTAATACCC 3476
QY      21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      3477 AGCTTAACCTGGCGGGAGCAAAACCAAGTGGAGGTGAGGTCCAGATTGTGCAACTGCT 3536
QY      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db      3537 GCCCAAACTCTCTGGCAACGTGCATCAATGGGGTGTCTGGACTGTCTACCAAGGGCC 3596
QY      61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      3597 GGAACGAGACCATCGCTCACCCCAAGGTCTCTGCATCCAGATGTATACCAATGTAGAC 3656
QY      81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      3657 CAAGACCTTGTGGCTGGCGCTCCGCAAGGTAGCGCTCATTCACACCTTGCACTTGC 3716
QY      101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db      3717 GGCTCTCGACCTTTACTGTGTACAGAGCACCGCATGTCAITCCGCTCGCGCGCGG 3776
QY      121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db      3777 GGTGATAGCAGGGCAGCTCTGCTGCCCGCGGCCATTTCTACTTGAAGGCTCTCTCG 3836
QY      141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db      3837 GGGGTCGCTGTGTGTCGCCCGCGGCACGCGCTGGGCATATTTAGGGCGCGGTGTGC 3896
QY      161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db      3897 ACCGTGGAGTGGCTTAGCGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGAACCAATG 3956
QY      181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db      3957 AGGTCCCGCGGTGTTACCGGATAACTCTCTCCACCAAGTAGTGCCTCCAGAGGCTTCAGG 4016

```

Combination of three hepatitis C virus antigens - used for detection of specific antibodies to diagnose infection.

Disclosure; Col 25-46; 57pp; English.

This sequence represents the Hepatitis C virus polypeptide coding sequence. Fragments of this sequence can be amplified and used in the combination of HCV antigens of the invention. The HCV antigen combination comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa) 1-120 of the HCV polypeptide), or its immunologically reactive fragment containing at least 8 aa. It also comprises two additional antigens from two different polypeptide domains, including at least 8 aa from the NS3, NS4, S or NS5 domains of the polypeptide, corresponding, respectively, to aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polypeptide. Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa regions of the HCV polypeptide. These antigen combinations are used diagnostically to detect anti-HCV antibodies, using any standard immunoassay format. These antigen combinations have a broader range of reactivity with antibodies than any antigen individually. (Updated on 25-MAR-2003 to correct PR field.)

Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

QY	281	SerGlyGlyAlaIatyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
DB	4257	TCGGGGGGCGCTTATGACATAAATAATTGTGACGAGTGCCACTCCACGGATGCCACATCC	4316
QY	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
DB	4317	ATCTTGGGCATTCGGCACTGCTCTTTGACCAAGACAGACTCGGGGGCGAGACTGGTTGTG	4376
QY	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
DB	4377	CTCGCCACCGCCACCCCTCCGGGCTCGTCACTGTGCCCCATCCCAACATCGAGGAGTT	4436
QY	341	AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle	360
DB	4437	GCTCTGTCCACCAACCGAGAGATCCCTTTTACGGCAAGGCTATCCCCCCTCGAAGTAAATC	4496
QY	361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla	380
DB	4497	AAGGGGGGAGACATCTCATCTTCTGTCATTCAAAGAAAGTCCGACGAACTCGCCGCA	4556
QY	381	LysLeuValAlaLeuGlyValAsnAlaValAlaIaTyrTyrArgGlyLeuAspValSerVal	400
DB	4557	AAGCTGTGTCCATTGGGCATCAATGCCGTGGCTACTACCCGGGTCTTGACGTGTCCGTC	4616
QY	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
DB	4617	ATCCCGACAGCGCGAGTGTTCGTCTGGCAACCGATCCCTCATGACCGGCTATACC	4676
QY	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
DB	4677	GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACCCAGACAGTCGATTCAGC	4736
QY	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
DB	4737	CTTGACCCCTACCTTCACATTGAGCAANTCACGCTCCCCAGGATGCTGCTCCCGCACT	4796
QY	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyLysIleTyrArgPheValAlaProGly	480
DB	4797	CAACGTCGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTGTGTGGCACCGGG	4856
QY	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
DB	4857	GAGCGCCCTCCGGCAGTTCGACTCGCTCGCTCTGTGAGTGCTATGACGCGAGGCTGT	4916
QY	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
DB	4917	GCTTGGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC	4976
QY	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
DB	4977	CCGGGGCTTCCCGTGTGCCAGGACCAATCTGAAATTTGGAGGGCGCTCTTTACAGGCCCTC	5036
QY	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
DB	5037	ACTCATATAGATGCCCATTTCTATCCGACAAAGCAGAGTGGGGAGAACCTTCCTTAC	5096
QY	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp	580
DB	5097	CTGTAGCGTACCAGCCACCGTGTGGGTAGGGCTCAAGCCCTCCCTCCCATCGTGGGAC	5156
QY	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
DB	5157	CAGATGGAGTGTGTGATTCGCTTCAGCCCCACCCCTCCATGGGCGCAACACCCCTGCTA	5216
QY	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
DB	5217	TACAGACTGGGGCGCTGTTCAGAATGAATAATCACCCCTGACGCACCCAGTCCACAAATCATC	5276
QY	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
DB	5277	ATGACATGCATGTCCGGCCACCTGGAGGTGCTCACGAGCACTGGGTGCTGTGTGGCGC	5336
QY	641	ValLeuAlaAlaLeuAlaIaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660

```

Db 5337 GTCTGCTGCTTTGGCGCGGTATGCTGCTCAACAGGCTGCTGCTAGTGGCGAGG 5396
Qy 661 lIeValleuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCTGCTTTGTCGGGAAGCGGCAATCATACCTGACAGGAAGTCTCTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAGAGTGC 5474
RESULT 14
AAD35043
ID AAD35043 standard; cDNA; 9401 BP.
XX AAD35043;
XX
XX 16-JUL-2002 (first entry)
XX Hepatitis C virus (HCV) polyprotein cDNA.
XX Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;
XX NS4 domain; S domain; NS5 domain; gene; ss.
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX CDS 342..9377
XX /*tag= a
XX /product= "HCV polyprotein"
XX /transl_except= (pos:366..368, aa:Xaa)
XX /note= "Xaa equals Lys or Arg"
XX /transl_except= (pos:372..374, aa:Xaa)
XX /note= "Xaa equals Thr or Asn"
XX /transl_except= (pos:867..869, aa:Xaa)
XX /note= "Xaa equals Thr or Ile"
XX /transl_except= (pos:1341..1343, aa:Xaa)
XX /note= "Xaa equals Val or Met"
XX /transl_except= (pos:2148..2150, aa:Xaa)
XX /note= "Xaa equals Ile or Leu"
XX /transl_except= (pos:2883..2885, aa:Xaa)
XX /note= "Xaa equals Asn or Tyr"
XX /transl_except= (pos:3681..3683, aa:Xaa)
XX /note= "Xaa equals Ser or Pro"
XX /transl_except= (pos:3690..3692, aa:Xaa)
XX /note= "Xaa equals Thr or Ser"
XX /transl_except= (pos:4167..4169, aa:Xaa)
XX /note= "Xaa equals Leu or Pro"
XX /transl_except= (pos:4323..4325, aa:Xaa)
XX /note= "Xaa equals Val or Gly"
XX /transl_except= (pos:4701..4703, aa:Xaa)
XX /note= "Xaa equals Tyr or Cys"
XX /transl_except= (pos:4752..4754, aa:Xaa)
XX /note= "Xaa equals Ser or Thr"
XX /transl_except= (pos:5970..5972, aa:Xaa)
XX /note= "Xaa equals Gly or Glu"
XX /transl_except= (pos:6183..6185, aa:Xaa)
XX /note= "Xaa equals His or Leu"
XX /transl_except= (pos:6186..6188, aa:Xaa)
XX /note= "Xaa equals Cys or Ser"
XX /transl_except= (pos:6402..6404, aa:Xaa)
XX /note= "Xaa equals Gly or Val"
XX /transl_except= (pos:7386..7388, aa:Xaa)
XX /note= "Xaa equals Ser or Thr"
XX /transl_except= (pos:7494..7496, aa:Xaa)
XX /note= "Xaa equals Phe or Tyr"
XX /transl_except= (pos:7497..7499, aa:Xaa)
XX /note= "Xaa equals Ala or Ser"
XX /transl_except= (pos:7845..7847, aa:Xaa)
XX /note= "Xaa equals Phe or Leu"
XX /transl_except= (pos:8409..8411, aa:Xaa)
XX /note= "Xaa equals Gly or Arg"
XX /transl_except= (pos:9102..9104, aa:Xaa)

```

```

FT /note= "Xaa equals Gly or Arg"
FT /transl_except= (pos:9327..9329, aa:Xaa)
FT /note= "Xaa equals pro or Leu; These translational
FT exceptions occur while decoding for the alternative
FT version of HCV polyprotein (AAE22052)"
XX
FN US6312889-B1.
XX
XX 06-NOV-2001.
XX
XX 12-MAY-1995; 95US-00440549.
XX
XX 04-APR-1990; 90US-00504352.
XX 07-JUL-1992; 92US-00910760.
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo Q, Kuo G;
XX
XX WPI; 2002-040268/05.
XX P-PSDB; AAE22049.
XX
XX Combination of hepatitis C viral (HCV) antigens, useful in improved
XX immunoassay for detecting HCV antibodies.
XX
XX Example 1; Fig 1; 58pp; English.
XX
XX The invention relates to combination of hepatitis C viral (HCV) antigens
XX that have a broader range of immunological reactivity than any single HCV
XX antigen. The combinations consist of an antigen from the C domain of the
XX HCV polyprotein, and at least one additional HCV antigen from either the
XX NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
XX the form of fusion protein, a simple physical mixture, or the individual
XX antigens commonly bound to a solid matrix. The combinations of antigens
XX provides broad range immunoassays for anti-HCV antibodies. The invention
XX therefore provides a method for detecting antibodies to HCV in a mammal
XX suspected of containing such antibodies. The present sequence is HCV
XX polyprotein cDNA
XX
XX Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.93e-241 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: Gaps: 0
US-09-930-591-2 (1-686) x AAD35043 (1-9401)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGGGCCCATCACGGCTAGCCGCCAGCAGACAGGGGCTCTTAGGTGCATAATCACC 3476
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThAla 40
Db 3477 AGCCTAACTGGCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3536
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db 3537 GCCCAAACTTCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGGCCC 3596
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAACGAGGACCATCGCTCAACCAAGGGTCTCTGCATCCAGATGTATACCAATGTAGAC 3656
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTTGGGGTGGCCCGCTCGGCAAGGTAGCCGCTCATTACACCCCTGCACCTTGC 3716
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120

```


Db 3717 GGCTCTCGACCTTTACCTGGTCACGAGGACGCCGATGCTATTCCTCGTGGCCGGCGG 3776
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuIysGlySerSer 140
Db 3777 GGTGATAGAGGGGACGCTGCTGTCGCCCGCGGCCAATTTCTTACTTGAAGGCTCCTCG 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGGTCCGTGTTGTGCCCGCGGGGCACGCCGTGGGCATATTTAGGGCCCGGGTGTGC 3896
Qy 161 ThrArgGlyValAlaIysAlaValAspPheIleProValGlnSerLeuGluThrThrMet 180
Db 3897 ACCGCTGGAGTGGCTAAGCGGGTGAATTTATCCCTGTGGAGAACCTAGACACAACCATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGGTCCCGGTGTTACGGGATACTCTCTCCACAGTAGTGCCCGAGGCTTCCAGGTG 4016
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyIysSerThrIysValProAlaIaTyrAla 220
Db 4017 GCTCACCTCCATGCTCCACAGCGAGCGGCAAGACCAAGGTCGCCGGTGCATATGCA 4076
Qy 221 AlaGlnGlyTyrIysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4077 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTCTGCTCAACACTGGGCTTTGGT 4136
Qy 241 AlaTyrMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4137 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 4196
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIysPheLeuAlaAspGlyIysCys 280
Db 4197 ACCACTGGACGCCCATCATCGTACTCACCCTAGCGCAAGTCTCTGCCAGCGGGGTGC 4256
Qy 281 SerGlyIysAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGGCGCTTATGACATAAATTTGTGACAGTGGCCACTCCACGGATGCCACATCC 4316
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCATCGGCATGTCCTTGTACCAAGACAGACTCGGGGGCGAGACTGGTGTGTG 4376
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4377 CTGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCATCCCAACATCGGAGGTT 4436
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyIysAlaIleProLeuGluAlaIle 360
Db 4437 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerIysIysIysCysAspGluLeuAlaIa 380
Db 4497 AAGGGGGGAGACATCTCATCTTCTGTCTCATTTCAAGAAGAAGTGGACGAATCGCCGCA 4556
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4557 AAGCTGGTGCATTTGGGCATCAATCGGTGGCTACTACCGGGTCTTGACGTGCCGTGC 4616
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4617 ATCCCGACACGCGCGCATGTTGCTGCGTGGCAACGATGCCCTCATGCCGGCTATACC 4676
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4677 GGGGACTTCGACTCGGTGATAGTACTGCATACGTGTGTCAACGACACAGTCCGATTCAGC 4736
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4737 CTTGACCTACCTTCACTACCTAGAGCAATCACGCTCCCGCAGGATGCTCTCTCCGCACT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyIysProGlyIleTyrArgPheValAlaProGly 480
Db 4797 CAACGTTCGGGGCAGGACTGGCAGGGGGAGACCCAGGCAATCTACAGATTTGTGGCACCGGG 4856

Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4857 GAGCCGCCCTCCGCGCATGTTGACTGTCGTCCTCTGAGTGCTATGAGCGAGGCTGT 4916
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGGTATGAGTCAACGCCCGCGAGACTACAGTTAGCTACGAGCGTACATGAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCGTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIysGlnSerGlyIysLeuAsnLeuProTyr 560
Db 5037 ACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerSerTrpAsp 580
Db 5097 CTGGTAGGCTACCAAGCCACCGTGTGCGCTAGGGCTCAAGGCCCTCCGCCCATCGTGGAC 5156
Qy 581 GlnMetTrpIysCysLeuIleArgLeuIysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCACCTCCATGGGCCAACCCCTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIysTyrIle 620
Db 5217 TACAGACTGGGCGCTGTTTCAAGATGAATCAACCTGACGCCACCCAGTCAACAAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrThrProValLeuValGlyGly 640
Db 5277 ATGACATCATGTCGGCGACCTGGAGTCTGTCAGAGCACTGGGTGCTCGTGGCGGC 5336
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCCGTGGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTCGGTGCTCATGTGGCAGG 5396
Qy 661 IleValLeuSerGlyIysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCGCTTTGTCCGGGAAGCCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAGAGTGC 5474
RESULT 15
ID AAN92103
ID AAN92103 standard; DNA; 6905 BP.
XX AAN92103;
AC AAN92103;
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
DE clones 12f through 15e.
XX Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH.
XX Hepatitis C virus.
OS
FH Key Location/Qualifiers
FT CDS 3..6905
FT /*tag= a
XX
PN EP318216-A.
XX
PD 31-MAY-1989.
XX
PF 18-NOV-1988; 88EP-00310922.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.

1693	ACC	CGT	GAG	TGG	CTA	AGG	CGG	TGG	AGC	ATT	TAT	CC	T	GTG	CGA	AAC	CT	TAG	ACA	CA	ACC	ATG	174
181	Arg	Ser	Pro	Val	Phe	Ser	Asp	Asn	Ser	Ser	Pro	Pro	Ala	Val	Pro	Gln	Ser	Tyr	Gln	Val	200		
1743	AGG	TCC	CGG	TGT	TCA	CGG	TAA	CT	C	T	C	T	C	CA	C	AG	TAG	T	CC	C	AG	AGT	1802
201	Ala	His	Leu	His	Ala	Pro	Thr	Gly	Ser	Gly	Leu	Ser	Thr	Leu	Val	Pro	Ala	Ala	Tyr	Ala	220		
1803	GCT	CAC	T	CTC	AT	G	T	C	C	C	A	G	C	A	G	G	G	A	A	A	G	1862	
221	Ala	Gln	Gly	Tyr	Leu	Val	Leu	Val	Leu	Asn	Pro	Ser	Val	Ala	Ala	Thr	Met	Gly	Phe	Gly	240		
1863	GCT	CAG	GCT	TAT	AAG	T	GCT	AGT	CTA	CA	C	C	C	C	T	CT	GT	CT	G	CA	1922		
241	Ala	Tyr	Met	Ser	Leu	Ala	His	Gly	Leu	Asp	Pro	Asn	Leu	Arg	Thr	Gly	Val	Arg	Thr	Leu	260		
1923	GCT	TAC	ATG	T	CCA	AGG	T	CGA	T	CGA	T	C	T	A	C	A	T	C	A	G	1982		
261	Thr	Thr	Gly	Ser	Pro	Leu	Thr	Tyr	Ser	Thr	Tyr	Gly	Leu	Ser	Phe	Leu	Ala	Asp	Gly	Cys	280		
1983	ACC	ACT	GGC	ACC	CC	ATC	A	G	T	ACT	CA	C	CT	ACG	C	CA	G	T	CT	C	2042		
281	Ser	Gly	Gly	Ala	Tyr	Asp	Leu	Leu	Leu	Cys	Asp	Glu	Cys	His	Ser	Thr	Asp	Ala	Thr	Ser	300		
2043	TGG	G	G	G	G	G	G	T	T	ATG	ACA	TAA	T	T	G	T	G	A	C	T	2102		
301	Leu	Leu	Gly	Leu	Gly	Thr	Val	Leu	Asp	Gln	Ala	Glu	Thr	Ala	Gly	Ala	Arg	Leu	Thr	Val	320		
2103	ATC	T	TGG	CAT	CGC	ACT	G	T	C	T	T	AC	CA	G	CA	G	ACT	C	G	G	2162		
321	Leu	Ala	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Val	Thr	Val	Pro	His	Pro	Asn	Leu	Glu	Val	340			
2163	CTG	CC	AC	CGG	CC	ACC	CT	C	CGG	GCT	CCG	T	CAC	T	G	T	CC	CC	CA	T	2222		
341	Ala	Leu	Ser	Thr	Thr	Gly	Glu	Leu	Pro	Phe	Tyr	Gly	Leu	Ala	Leu	Pro	Leu	Glu	Ala	Leu	360		
2223	GCT	CT	G	T	CT	CC	AC	C	AC	CG	AG	A	G	A	T	CC	C	C	T	C	2282		
361	Lys	Gly	Gly	Arg	His	Leu	Leu	Phe	Cys	His	Ser	Lys	Lys	Lys	Cys	Asp	Glu	Leu	Ala	Ala	380		
2283	AGG	G	G	G	G	G	A	CA	T	C	T	C	T	CT	CA	T	CA	A	A	A	2342		
381	Lys	Leu	Val	Ala	Leu	Gly	Val	Asn	Ala	Val	Ala	Tyr	Tyr	Arg	Gly	Leu	Asp	Val	Ser	Val	400		
2343	AAG	CTG	G	TCG	CA	T	TGG	GC	AT	CA	A	T	G	CG	T	ACT	AC	C	G	G	2402		
401	Leu	Pro	Thr	Ser	Gly	Asp	Val	Val	Ala	Ala	Thr	Asp	Ala	Leu	Met	Thr	Gly	Phe	Thr	420			
2403	AT	C	CG	CA	C	AC	G	CG	CA	T	G	T	G	C	A	C	G	A	T	2462			
421	Gly	Asp	Phe	Asp	Ser	Val	Leu	Asp	Cys	Asn	Thr	Cys	Val	Thr	Gln	Thr	Val	Asp	Phe	Ser	440		
2463	GCG	CA	T	T	CG	AT	T	G	CA	T	G	CA	A	T	AC	T	GT	G	T	2522			
441	Leu	Asp	Pro	Thr	Phe	Thr	Leu	Glu	Thr	Leu	Thr	Leu	Pro	Gln	Asp	Ala	Val	Ser	Arg	Thr	460		
2523	CTT	G	AC	CT	T	AC	CA	T	T	G	A	C	A	T	C	A	C						

```
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIysGlnSerGlyGluAsnLeuProTyr 560
Db |||||
Db 2823 ACTCATATAGATGCGCCACTTCTATCCACACAAGCAGAGTGGGGAGAACCTTCCTTAC 2882
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db |||||
Db 2883 CTGGTAGCGTACCAAGCCACCGTGTGGCTAGGGCTCAAGCCCTCCCCCATCGTGGGAC 2942
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db |||||
Db 2943 CAGATGTGGAGGTGTTGATTCGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 3002
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db |||||
Db 3003 TACAGACTGGGCGCTGTTCAAGATGAATCACCCTGACGACCCAGTCACCAATACATC 3062
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGly 640
Db |||||
Db 3063 ATGACATGCATGTGCGCGCACCTGGAGTCTCACGAGCACCTGGGTGCTCTGTGGCGGC 3122
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db |||||
Db 3123 GTCCTGGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTGCGTGTATAGTGGCAGG 3182
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db |||||
Db 3183 GTCGTCTGTTCGGGAGACCGGCAATCATCTACAGGGAGTCTCTACCGAGAGTTC 3242
Qy 681 AspGluMetGluGluCys 686
Db |||||
Db 3243 GATGAGATGGAAGAGTGC 3260
```

Search completed: September 16, 2004, 00:57:45
Job time : 913 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 15, 2004, 15:46:56 ; Search time 4559 Seconds
(without alignments)
4493.410 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPITAYAQOTRGLGCIIT.....PALIPDREVLVREFDEMEEC 686

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DRV=xlh
-Q=/cgn2_1/USPTO_spool/US0930591/runat_15092004.164646.7108/app.query.fasta_1.839
-DB=EST -Qfmt=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0930591 @CGN 1.1 3437 @runat_15092004.164646.7108 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estnu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_htc:
9: gb_est1:
10: gb_est2:
11: gb_htc:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estcm:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pln:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_rod:
26: em_gss_phg:
27: em_gss_vrl:
28: gb_gss1:

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	163	4.5	3291	11	BC046781	BC046781 Mus muscu
2	150.5	4.2	3229	11	AK004665	AK004665 Mus muscu
3	141	3.9	3359	29	AY411078	AY411078 Homo sapi
4	140.5	3.9	822	13	BU054791	BU054791 UI-M-PD0-
5	136.5	3.8	3320	11	BC059369	BC059369 Homo sapi
6	133.5	3.7	889	14	CD359697	CD359697 AGENCOURT
7	132.5	3.7	818	13	BU187274	BU187274 AGENCOURT
8	131	3.6	1201	9	AL560974	AL560974 AL560974
9	129	3.6	2972	11	AK031534	AK031534 Mus muscu
10	129	3.6	2984	11	AK031679	AK031679 Mus muscu
11	129	3.6	2986	11	AK028274	AK028274 Mus muscu
12	129	3.6	3956	11	BC043699	BC043699 Mus muscu
13	128.5	3.6	1697	29	AY404177	AY404177 Mus muscu
14	127.5	3.5	691	13	BU054966	BU054966 UI-M-PD0-
15	127.5	3.5	738	14	CD240900	CD240900 AGENCOURT
16	127.5	3.5	2976	29	AY400284	AY400284 Homo sapi
17	126	3.5	1283	13	BQ709745	BQ709745 AGENCOURT
18	124.5	3.4	623	14	CF131593	CF131593 UI-HP-FQ0
19	124.5	3.4	2388	29	AY418898	AY418898 Mus muscu
20	124.5	3.4	3633	11	AY383690	AY383690 Rattus no
21	124	3.4	2388	29	AY418896	AY418896 Homo sapi
22	123.5	3.4	1788	11	AY105041	AY105041 Zea mays
23	123	3.4	790	14	CK017540	CK017540 AGENCOURT
24	123	3.4	3461	29	AY398774	AY398774 Mus muscu
25	123	3.4	3802	11	AK004733	AK004733 Mus muscu
26	123	3.4	4327	11	BC058331	BC058331 Mus muscu
27	123	3.4	4640	11	BC062885	BC062885 Mus muscu
28	122	3.4	3879	11	BC028405	BC028405 Homo sapi
29	121.5	3.4	3211	11	BC041392	BC041392 Homo sapi
30	121	3.3	2270	29	AY418897	AY418897 Pan trogl
31	120.5	3.3	478	10	BF725559	BF725559 bx18a02.Y
32	120.5	3.3	694	12	BI088407	BI088407 602852901
33	120.5	3.3	3856	11	AK078552	AK078552 Mus muscu
34	120	3.3	1143	12	BM926541	BM926541 AGENCOURT
35	120	3.3	3494	11	BC013208	BC013208 Homo sapi
36	119	3.3	919	13	EX424517	EX424517 BX424517
37	119	3.3	1339	14	CF753594	CF753594 EST-Conti
38	119	3.3	3679	11	AK084541	AK084541 Mus muscu
39	118.5	3.3	769	9	AU170397	AU170397 AU170397
40	118.5	3.3	1008	12	BQ050646	BQ050646 AGENCOURT
41	117.5	3.2	2031	11	BC038434	BC038434 Homo sapi
42	117.5	3.2	3024	11	AK049321	AK049321 Mus muscu
43	117	3.2	3008	11	AF318356	AF318356 Homo sapi
44	117	3.2	3461	29	AY398772	AY398772 Homo sapi
45	117	3.2	3915	11	BC038417	BC038417 Homo sapi

ALIGNMENTS

RESULT 1
BC046781
LOCUS BC046781
DEFINITION Mus musculus DEAH (Asp-Glu-Ala-His) box polypeptide 16, mRNA (CDNA clone IMAGE:5716182), containing frame-shift errors.
ACCESSION BC046781
VERSION BC046781.1
KEYWORDS HTc.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3291)

US-09-930-591-2 (1-686) x BC046781 (1-3291)

QY	89	ProGlnGlyAlaArgSerLeuThrPro-----CysThrCys-----	100
Db	953	CCGAGAGCAGGAGAAGCTGGAGGCCACTAAATCGCTACCACATGCCCAAGGACACCCGAG	1012
QY	101	GlySerSer---AspLeuTyrLeuValThrArgHisAlaAspValIleProValatArg	119
Db	1013	GGCAGCCAGCTGGAATGTGGACATCGTTCAGGAGAAATCGGTGCGCCGGGAGAGGAGC	1072
QY	120	ArgGlyAspGlyArg-Gly-----SerLeuLeuSerProArgpr	132
Db	1073	AGCGGCGCTGGGAGGAGGCCAGCTGGGTGCTGCATCCCTGAATTTGGAGCCCGAGATG	1132
QY	132	olIleSerTyrLeuLysGlySerSerGlyProLeuLeuCysProAlaGly-----	149
Db	1133	CTGC-----TGGCAGAGGAGCCAAATACCAGCTGTGTCTGGAGGAGGAGC	1177
QY	150	----HisAlaVal-----GlyIlePheArgAla	158
Db	1178	AGACATTCAGTTTGTTCGTGCTCACTCCAGGTCACAGAGGCGTCAGGCCCGC	1237
QY	158	aValCysThrArgGlyValAla-LysAlaValAspPheLeProValGluSerLeuGlu	178
Db	1238	CCCTGTCCAGCCAGGCCAGCAGCAAG-----GAATCATCCAG	1276
QY	178	hrThrMetArgSer---ProValPheSerAspAsnSerSerProProAlaValProGln	197
Db	1277	CTGTGGCCGCGAGCCGCGCGTGTTCCTTCGAGAGGAGCTCTGGCCGCCATTGCCA	1336
QY	197	erTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValPro-	216
Db	1337	ACCATCAGTCTCATCATCGAAGGCGAGACTGGCTCTGGGAAGACCACACAGATCCCAC	1396
QY	217	-----AlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	232
Db	1397	AGTACCTCTTTGAGGAGGGTTTACAAAGAAGGGCATGAAGATTGCTCCACCCAGCC	1456
QY	232	erValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisGly-----	248
Db	1457	GGAGAGTGGCGGTATGACTGTGGAGCCCGAGTGGCCCGGAGATGGGTGTGAAGCTTG	1516
QY	249	-----IleAspProAsnIleArg-----ThrGlyValArgThrIleThr	262
Db	1517	GGAAAGAGGTGGGTACAGCATCCGGTTTGAGAGCTGCACCTCAGAGCAACTGTGT---	1572
QY	262	hrGlySerProIleThrTyrSerThrTyrGly-----LysPheLeuAlaAspG	278
Db	1573	-----CTCGCTATATGACAGATGGAAATGCTACTCCGAGAGTTTCTCTCTGAGC	1621
QY	278	lyGlyCysSerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspA	298
Db	1622	CTGACCTTGCA-----AGTTACAGTGTGTCATGTCATGAAGCTCAGAGCGGACCT.	1675
QY	298	laThrSerIleGluGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArg	318
Db	1676	TGCACACAGACATCTCTTTGATTGATCAAAAGCTCGCTAGATTCCGACCTGAGCTCA	1735
QY	318	euThrValLeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIle	338
Db	1736	AGGTCTGTGTCTTCAGCCACA-----	1758
QY	338	luGluValAlaLeuSerThrThrGlyIleProPheTyrGlyIysAlaIleProLeuG	358
Db	1759	-----CTGATATGCGCGGTTTCTGTCTTCTTCATCAGCGC-----CTGTCT	1804
QY	358	luAlaIleIysGlyIleArgHisLeu-----IlePheCysHisSerLysLysCysA	376
Db	1805	TCAGAAATCCCTGGACGAGGTTTCCAGTTGACATCTTATACCAAGGCCCCAGAGGCTG	1864
QY	376	spGluLeuAlaAlaLysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGly	396
Db	1865	ACTACTCGAAGCCTCGGTTGTGTCTGTG-----	1894

```

email: ccsp@cs-remail.nlm.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
Thomas L. Casavant.
Web site: http://genome.uiowa.edu
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: frame shifted.

FEATURES
    source
        1..3291
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone="IMAGE:5716182"
            /tissue_type="Mouse, brain 12.5 dpc"
            /clone_lib="NIH EMAP_FCO"
            /lab_host="DH10B"
            /note="Vector: pYX-ASC"

ALIGNMENT SCORES:
    Alignment Score: 0.00271
    Score: 163.00
    Percent Similarity: 35.15%
    Percent Local Similarity: 23.87%
    Query Match: 4.51%
    DB: 11
    Length: 3291
    Matches: 127
    Conservative: 60
    Mismatches: 165
    Indels: 181
    Gaps: 27

```


Percent Similarity: 34.22%
 Best Local Similarity: 21.45%
 Query Match: 4.16%
 DB: 11
 Conservatives: 72
 Mismatches: 221
 Indels: 151
 Gaps: 24

US-09-930-591-2 (1-686) x AK004665 (1-3229)

126 SerLeuLeuSerPro---ArgProIleSerTyLeuLysGlySerSerGlyProLeu 144
 6 AGCATGGCTGGCCCTGGGACCAAGTAAATCTGGAACCGGGTACAGAAAGGCGCA--- 62
 145 LeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyVal 164
 63 -----GCTGTCACATTTCTGAAGAGAGACAAAGTCTC---ACCGAAGACTCT 107
 165 AlaLysAlaValAspPheIleProValGluSerLeu-----GluThrThrMetArg 181
 108 GGGACAACCATCATTTACAATCCCTACGCTGCCCTTTCTATAGCAGCAGCGGCGAGAAG 167
 182 SerProValPheSerAspAsnSerSerProAlaValProGlnSerTyGlnValAla 201
 168 CTGCCCGCTGTTCAAGCTTAGAACCACTTTTGTACTTTGTAGAAACTATCAGACCGTG 227
 202 HisLeuHisAlaProThrGlySerGlyLysSerThrLysValPro----- 216
 228 GTGATTGTTGGAAACAGCATGTGGAGAGACCCAGATCCACAGTACCTGGCGAGAA 287
 217 AlaAlaTyAlaAlaGlnGlyTyLysValLeuValLeuAsnProSerValAlaAlaThr 236
 288 GCTGGCTGGACAGCGAAGCAGGAGTGTCTGGAGTGCACCGCTCGCAGAGTGGCTGCC 347
 237 MetGlyPheGlyAlaTyMetSerLysAlaHisGlyIleAspProAsnIleArgThrGly 256
 348 GTGACGCGATGA-TCTTTCCTCCCAAGCTTGCAGGCGGAGTGTGATGAAGGGGTGCA 406
 257 ValArgThrIleThrGlySerProIleThrTy-----Ser 269
 407 GTGTGGCGCATGAAGTGGGCTACTGTATCCGCTTGTGATGACTGCACCGACCCACTGGCC 466
 270 ThrTyGlyLysPheLeuAlaAspGlyGlyCysSerGly----- 282
 467 ACCAGAAATATAGTGTCAATCATGCTGTGATGAGTGTCTGGTCAGAGAAATGATGGTGCCTG 526
 283 ---GlyAlaTyAspIleIleCysAspGlyCysHisSerThrAspAlaThrSerIle 301
 527 TTAACAAATATAGTGTCAATCATGCTGTGATGAGTGTCTGGTCAGAGAAATGATGGTGCCTG 586
 302 LeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeu 321
 587 ATTGCCATTGGCTGTGTAAGAAGATTCAGAAAAGCGAGGGGATCTTCGCTTGATTGTG 646
 322 AlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGlu----- 339
 647 GCCTCAGCCACTCGGACGCGAGAGAAATTCAGATTTCCTTAACAGAGATGAGACCACT 706
 340 -----ValAlaLeuSerThrThrGlyGlu----- 347
 707 GACCCAGCCAGAGATACCTCTGTGACGCTCAGATGGAAGCAGGACATTCACAGTGGAT 766
 348 -----IleProPheTyGlyLysAlaIleProLeuGluAlaIle 360
 767 ATCTTTTACACAAAGTCTGTTCAGATATATCAGGCACTGTGGACACCGTGGTA 826
 361 Lys-----GlyGlyArgHisLeuPheCysHisSerLysLysLys 374
 827 AAAATTTCATCAGACAGAGAGATGAGACATACCTAGCTTTCTTACTGCCAGGAGAA 886
 375 CysAspGluLeuAlaAlaLysLeuVal-----AlaLeuGly----- 386
 887 GTAGAGACTGTGTGTCCATGTTGATCAGACGCGCGCGCTGGCTCGCACTGGGATG 946
 387 -----ValAsnAlaAlaAlaTyTyArgGlyLeuAsp----- 397

947 AAGAAACACCTCCGGTCTCTCCCATGTATGACGAGTGCCTTCTTTTGACGAGATGAAG 1006
 398 ---ValSerValIleProThrSerGlyAspValValValValAlaThrAspAlaLeuMet 416
 1007 GTGTTTGAAGGGTGTACACAGCGTCAAGGAAGTATCGTGCCCAATATGTGGCAGAG 1066
 417 ThrGlyPheThr---GlyAspPheAspSerValIleAspCysAsnThrCysValThrGln 435
 1067 ACTTCCATCAACATCACTGGCATGTGTATGTATGATGATGCTGTGCTTATGAAGCTCGGA 1126
 436 ThrValAspPheSerLeuAspProThrPheThrIleGlu---ThrIleThrLeuProGln 454
 1127 -----GCTACAAACCCAGCAGACACTATTCAATGCTTGTGTGGTGTACCAAGTG 1174
 455 AspalavalSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGly---Ile 473
 1175 TCTCAGCGTCAAGCAATCAGCGGAGGAGCGTGTGGCGCAACCGCTCGGAAGTGT 1234
 474 TyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSerValLeuCys 493
 1235 TATCGCTTTACACA-----GAGGAAGCCTTTGACCACTACCTCAGTCCACCGTCCCT 1288
 494 GluCysTyraAlaGlyCysAlaTyTyTyGluLeuThrProAlaGluThrValArg 513
 1289 GAGATGCGAGCGACCAATTG-----GCCCGCTCATCTGTCAG 1327
 514 LeuArgAlaTyMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPheTrp 533
 1328 CTAAGGCC----- 1336
 534 GluGlyValPheThrGlyLeuThrHisIle---AspAlaHisPheLeuSerGlnThrLys 552
 1337 -----CTAGGATAGACATGCTCAGGTTCACCTTCATGCT----- 1375
 553 GlnSerGlyGluAsnLeuProTyLeuValAlaTyGlnAlaThrValCysAlaArgAla 572
 1375 ----- 1375
 573 GlnAlaProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLysProThr 592
 1376 -----CCCGCGCAGCA---CAGTCGATGTTCAAGCTTG----- 1408
 593 LeuHisGlyProThrProLeuLeuTyArgLeuGlyAlaValGlnAsnGluValThrLeu 612
 1409 -----GAGCTGCTCTATGCTCTCGAGGTCTGGACAAAGACTGCGCTA 1453
 613 ThrHisProVal 616
 1454 ACTGAGCTCTT 1465
 RESULT 3
 AY411078 3359 bp DNA linear GSS 16-DEC-2003
 LOCUS Homo sapiens TIE gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION Genomic survey sequence.
 ACCESSION AY411078
 VERSION AY411078.1 GI:39767046
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3359)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3359)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,

Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

TITLE

JOURNAL Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES

source

1..3359

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..3359

/gene="TIE"

/locus_tag="HCM4104"

ORIGIN

Alignment Scores:

Pred. No.: 0.323 Length: 3359
 Score: 141.00 Matches: 168
 Percent Similarity: 31.38% Conservative: 84
 Best Local Similarity: 20.92% Mismatches: 233
 Query Match: 3.90% Indels: 318
 DB: 29 Gaps: 41

US-09-930-591-2 (1-686) x AY411078 (1-3359)

QY 39 ThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHis 58
 DB 42 ACGAGCCCCAGCGCTTCTTCGTGACTTGGGTCTGGGAGGCC----- 86
 QY 59 GlyAlaGly-----ThrArg 63
 DB 87 GGGCGGGAGGGGCTCGGACGCTGGGCGCGCCCTGCTGCTGGAGAGACGACCT 146
 QY 64 ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValaspGlnasp--- 82
 DB 147 ATCGTCGCACCCCGCGCGCCACCCCTCGCTCGCGCGCAACAGCGTTTCGCACACGATC 206
 QY 83 ---LeuValGlyTrp-ProAlaPro----- 89
 DB 207 ACGCTTCGGGCTTCTCCAGCCCTCGGACCTCGTGGGCGTCTTCTCTCGTGGCGGT 266
 QY 90 ---GlnGlyAlaArgSerLeuThrProCysThrCysGlySerAspLeuTyrLe 107
 DB 267 GCTGGGGCGCGCGCGCGCTACCTAGCTGACACAGCGCCCTGGAGCCACCTGCTT 326
 QY 107 uValThrArgHis-----AlaaspValIleProVal----- 117
 DB 327 CCAGACAGGTCAACACACTGTGAACAAAGGTGACCGCTGACTTCTTTCGCAGGTGTG 386
 QY 118 -----ArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProil 133
 DB 387 CACAAGGAGAGCAGACAGCGTGTCTGGAAGAGCAACGGATCTTACTTCTACACCTG 446
 QY 133 eSerTyrLeuLysGlySerSerGlyCysProLeuLeuCysProAlaGlyHisAlaValGl 153
 DB 447 GACTGGCATGAAGCCAGGATGGGCGGTTC-----TCTGCGAGCTCCCAATGTGCAG 500
 QY 153 yIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheileProVa 173
 DB 501 CCACCATCGAGCGGCA-----TCTACAGTGCCTACT 530
 QY 173 lGluSerLeuGluThrThrMetArgSerProValPheSerAspAsn----- 188
 DB 531 TACCTGGAAAGCCAGCCCTCGGACGCGCTCTTTCGGCTCATCGTGGCGGTTGTGGG 590
 QY 189 -----SerSerProProAlaValProGlnSerTyrGlnValAlaHisLeu----- 203
 DB 591 GCTGGCGGTGGGGCCAGCTGTACCAAGAGTGCACAGGTGCTCATCATGAGAGGTGC 650
 QY 204 -----HisAlaProThrGlySer---GlyLysSerThrly 214

DB 651 TGCCACGACCATGACGGGAAATGTGTATGCCCCCTGGCTTCACTGGCACCCGCTGTGAA 710
 QY 214 sValProAlaAlaTyrAla-----AlaGlnGlyTyrLy 225
 DB 711 CAGGCTGTCAGAGAGGGCGGTTTGGCAGAGCTGCCAGGACGATGCCAGGCATATCA 770
 QY 225 sValLeuValLeuAsnProSer-----ValAlaAlaThrMetGlyPheGl 240
 DB 771 GGCTCCCGGGCCTCACCTTCTGCTCCAGACCCCTATGGCTCTTGTGGATCTGCC 830
 QY 240 yAlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaThrI 260
 DB 831 TGGAGAGGAGCCAGTGCACAGACCTGTGCCCCCTGTCATTGTTGGG----- 879
 QY 260 eThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGly 280
 DB 879 ----- 879
 QY 280 sSerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSe 300
 DB 880 -----CTGATTGCCAGCTCCAGTCCAGTGTGCAGAAATGGTGGC 917
 QY 300 rIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVa 320
 DB 918 ACTTGTG---ACCGGTTCAAGTGTGTCTGCTGCCCTCTGGGTGCGCATGGAGTGCACGT 974
 QY 320 lLeuAla---ThrAlaThrPro----- 326
 DB 975 GAGAACTCAGACCGGATCCCCAGATCTCAACATGGCTCAGAACTGGAGTTCAACTTA 1034
 QY 327 -----ProGlySerValThr-----ValProHisPro----- 335
 DB 1035 GAGACGATGCCCGGATCACTGTGACGTGCAGGGAACCCCTTCCCGTGGGGGACG 1094
 QY 336 -----AsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyLy 353
 DB 1095 ATAGAGCTACGAAAGCCAGACGCGACTGTCTCTGTCCACCAAGGCCATTGTGAGCA 1154
 QY 353 sAlaIlePro-----LeuGluAlaIleLysGlyGl 363
 DB 1155 GAGAAACACAGCTGAGTTCGAGGTGCCCGCTTGGTTCTTGGGACAGTGGGTCTTG 1214
 QY 363 yArgHisIleuPheCysHisSerLysLysCysAspGluLeuAlaLysLeuVa 383
 DB 1215 CAGTGGCGGTG-----TGTCACATCTGGCGGCCAAGACAGC 1250
 QY 383 l-AlaLeu-GlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
 DB 1251 CGGCGCTTCAAGGTCAAT-----GTGAAAGTCCCGCC 1283
 QY 403 ThrSerGlyAspValValAlaThrAspAlaLeuMetThr----- 417
 DB 1284 -----GTGCCCTGCTGCACCTCGGCTCTGACCAAGCAGACGCGCCAG 1328
 QY 418 -----GlyPheThrGlyAsp----- 422
 DB 1329 CTTGTGTCTCCCGCTGGTCTGCTCTCTGGGATGGACCCATCTCCACTGTCGCGCTG 1388
 QY 423 -----PheAspSerValIleAspCysAsnThrCysValThrGlnThrValAsp 438
 DB 1389 CACTACCGGCCCCAGGACAGTACCATGACCTGGTTCGACCATTTGTG----- 1433
 QY 439 PheSerLeuAspProThrPheThrIleGluThrIleThrLeu-----ProGln 454
 DB 1434 -----GTGACCCCACT-----GAGAACTGACGTAAATGAACCTGAGGCAAG 1478
 QY 455 AspAlaValSerArgThrGlnArgGlyArgThrGlyArgGlyLysProGlyIleTyr 474
 DB 1479 ACAGATACAGTGTCTGTGAGCTGAGCCGCGCCAGGGAAGGA----- 1523
 QY 475 ArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSer-----SerVa 491


```

Db 453 -----CTCCGCTACATGACAGATGGATGCTACTCCGAGAGTTCCTCTCTGAG 500
QY 278 GlyGlyCysSerGlyGlyAlaTyrAspIlelleCysAspGluCysHisSerThrAsp 297
Db 501 CTGAGACTTGCA-----AGTTACAGTGTGTGATGGTGAAGCTCAGCAGCGGACC 554
QY 298 AlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArg 317
Db 555 TTCCACACAGACATTCTCTTGGATTGATCAAGACGTGCTAGATCCGACTGAGCTC 614
QY 318 LeuThrValLeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIle 337
Db 615 AAGGCTCTGTGCTTCCAGCACA----- 638
QY 338 GluGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyIysAlaIleProLeu 357
Db 639 -----CTGATATGTCGCCCGTTTTCGCCCTTTCGATGACCCCTGCTTCA 686
QY 358 GluAlaIleGlyGlyGlyArgHisLeu--IlePheCysHisSerIysIysCysAspG 377
Db 687 GAATCCTGGACGACGAGTTTCCAGTTGACATCTTCTATACCAAGGCCCGAGGCTGACT 746
QY 377 lueuAlaIalysLeuValAlaLeuGlyValAsnAlaValaTyrTyrArgGlyLeuA 397
Db 747 ACCCTGGAGCCTCGTTGTGTGTG-----CTGC 776
QY 397 spValSerValIleProThrSerGlyAspValValVal 409
Db 777 AGATTACAGTGCACCCGCCCTCGAGATATACTGTG 814

RESULT 5
BC059369
LOCUS
DEFINITION
Homo sapiens cDNA clone IMAGE:30346256, containing frame-shift errors.
ACCESSION
BC059369
VERSION
BC059369.1 GI:37590735
KEYWORDS
HTC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3320)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carrincci,P., Frange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Hellon,E., Kertanen,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blackley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 15899-15903 (2002)
22388257
12477932
PUBLISHED
2 (bases 1 to 3320)
Strausberg,R.
Direct Submission
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgsbbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 133 Row: 1 Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 20544128
This clone has the following problem: frame shifted.

FEATURES
Location/Qualifiers
1..3320
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30346256"
/tissue_type="placenta, normal"
/clone_lib="NIH_MGC_147"
/lab_host="DH10B"
/note="vector: pluescript"

ORIGIN
Alignment Scores:
Pred. No.: 0.836 Length: 3320
Score: 136.50 Matches: 114
Percent Similarity: 33.69% Conservative: 73
Best Local Similarity: 20.54% Mismatches: 216
Query Match: 3.77% Indels: 153
DB: 11 Gaps: 24

US-09-930-591-2 (1-686) x BC059369 (1-3320)
QY 132 ProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAla 151
Db 35 CCGTGAAGTTCTCGCCAGCCCGGTACAGAGGGGCCA-----GGTGAAGC 79
QY 152 ValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheile 171
Db 80 ATCTCTGAAGAGACACAAAGTCTGGCTGAAAACCTCT---GGGACACGGTGTGTACAC 136
QY 172 ProValGluSerLeu-----GluThrThrMetArgSerProValPheSerAspAsn 188
Db 137 CTTATGCTGCCCTTTCATAGACGACAGGAGGAGAGCTGCCGGTATTCAAGCTTAGG 196
QY 189 SerSerProProAlaValProGlnSerTyrGlnValAlaHisLeuHisAlaProThrGly 208
Db 197 AATCATATTTTATCTGTATAGAAAATATTCAGACGGTGTGTGTGTGGTGAACAGGA 256
QY 209 SerGlyLysSerThrLysValPro-----AlaAlaTyrAlaAlaGlnGly 223
Db 257 TGTGGGAAGAGCACACAGATTCTCTCAGTACCTTCAGAAAGCCGGCTGGACAGCTGAAGGA 316
QY 224 TyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGlyAlaTyrMet 243
Db 317 AGATGTGTAGAGTGAACCCAGCTCGTGAAGAGTGGCTGCTGTACAGTTGCGGAGGAGTA 376
QY 244 SerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIleThrThrGly 263
Db 377 GCTCAAGAAAGGGGTGC-AGTGTGGGCCACAGAGTGGGCTACTG----- 420
QY 264 SerProIle-----ThrTyrSerThrTyrGlyLysPheLeuAlaAspGly 278
Db 421 CATCCGCTTTGATGACTGCACCGCAGCTGGCCACGAGAAATTAAGTTTCTTACTGATGGA 480

```

279 GlyCysSerGly-----GlyAlaThrAspIleIleIleCys 290
 Db ATGCTGGTCAGGAATGATGTTGATCGCTGTTTAAACAAATATAGTGCATCATGCTG 540
 291 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln 310
 Db GATCAGACCCACGAGAGGACCTTGTACACTGACATGCAATGGCTTGTCTAAAAAGATT 600
 311 AlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThrAlaThrProGlySerVal 330
 Db CAGAAAAGCGAGGAGATCTTCGATTGATGTTAGCTTACGCCACTCTGGATGCAGACAAA 660
 331 ThrValProHisProAsnIleGluGlu-----ValAla 341
 Db TTCGGGATTTCTTTAATCAAAATGAACCCAGTGATCCAGCAGGATACATGTGTGATC 720
 342 LeuSerThrThrGlyGlu-----IlePro 349
 Db CTTACAGTGAAGGAGGAGAACATTTCCGGTGGATATCTTTATACAAAGTCTGTCTCA 780
 350 PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyGly 363
 Db GATTATATCAATCACTGTCGAACCTGTGTGTGAATTCACACAGACAGGAGACGCA 840
 364 ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaIleVal 383
 Db GAGCTTTTAGCATTTCTTACTGCGCAGGAGAGGTAGAACTGTGTGTGATGCTCATC 900
 384 -----AlaLeuGly-----ValAsnAlaValAlaIleTyr 392
 Db GAGCAGCTCGACACTAGTCTGCATCGCATGGAAGACACCTCCGAGTCTCCCCATG 960
 393 TyrArgGlyLeuAsp-----ValSerValIleProThrSerGly 405
 Db TATCGAGGACTGCTCTCTTGTGCAAAATGAAAGTGTGTGAAAGGTTGTCACGAGTGC 1020
 406 AspValValValAlaThrAspAlaLeuMetThrGlyPheThr---GlyAspPheAsp 424
 Db AGAAGGTGATAGTGGCCACCAATGTGGCAGAAACCTCTATCAATCAATGCGGCATG 1080
 425 SerValIleAspCysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThr 444
 Db TATGTGATCGACTGTGCTTTGTGAAACTCCGA-----GCTTACAAATCCAGG 1128
 445 PheThrIleGlu---ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArg 463
 Db ACAGCTATTGAATGCTTTGGTGGTGGTCCAGTCTCCCAAGCATCACTAATCAGCGAGCA 1188
 464 GlyArgThrGlyArgGlyLysProGly---IleTyrArgPheValAlaProGlyGluArg 482
 Db GGAGTGGTGTGCTAGTCTGCGGAAATGTTATCGCTTTATACA-----GAGGAA 1242
 483 ProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrp 502
 Db GCCTTTGACAAGTTGCCCTCATGCTACGTTCTCGAGATGCGCGTAGTAATTG----- 1296
 503 TyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThrProGly 522
 Db -----GCACTGTCTATCTCGACTGAAAGCA----- 1323
 523 LeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHis 542
 Db -----CTAGAAATTGCAAT 1338
 543 Ile---AspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeu 561
 Db GTCTCAGGTTCCACTTCATGTCG----- 1362
 562 ValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAspGln 581
 Db -----CCCCCTCCAGCA-----CAGTCG 1380

582 MetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeuTyr 601
 Db ATGTTTCAAGCTTGT-----GAGTTACTGTAT 1407
 602 ArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProVal 616
 Db GCTCTGGAGGCTCTGGCAAGAACTGTGCGCTAACTGAACCGCTT 1452
 RESULT 6
 LOCUS CD359697 889 bp mRNA linear EST 29-MAY-2003
 DEFINITION AGENCOURT 14284682 NIH_MGC_180 Homo sapiens cDNA clone
 IMAGE:30388971 5', mRNA sequence.
 ACCESSION CD359697
 VERSION CD359697.1 GI:31131108
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (Bases 1 to 889)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM461 row: 1 column: 04
 High quality sequence start: 22
 High quality sequence stop: 653.
 FEATURES
 Location/Qualifiers
 1..889
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30388971"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_lib="NIH_MGC_180"
 /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site: 1. Not I;
 Site_2: EcoRV (destroyed); Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.68 kb. Library was
 constructed by (Invitrogen). Note: this is a NIH_MGC
 Library."
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.157 Length: 889
 Score: 133.50 Matches: 61
 Percent Similarity: 40.24% Conservative: 40
 Best Local Similarity: 24.30% Mismatches: 95
 Query Match: 3.69% Indels: 55
 DB: 14 Gaps: 11
 US-09-930-591-2 (1-686) x CD359697 (1-889)
 183 ProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnValAlaHis 202
 Db CCGGTGTTCCATTCGAGAGGAGCTCTGCTGCTATTGCAATCACCAGTCTCTCATC 103
 203 LeuHisAlaProThrGlySerGlySerThrLysValPro-----Ala 217
 Db ATTGAAGCGGAGAGAGGCTCAGGGAAGACACCAGATCCCGCAGTATCTCTTTGAGGAG 163

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

REFERENCE

6 (bases 1 to 2972)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furumoto, N., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

REFERENCE

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

COMMENT

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://phantom.gsc.riken.go.jp/

FEATURES

Location/Qualifiers
1..2972
/organism="Mus musculus"
/mol_type="mRNA"

source

/strain="C57BL/6J"
/db_xref="FANTOM_DB:6030448F20"
/db_xref="MGI:2394772"
/db_xref="taxon:10090"
/clone="6030448F20"
/sex="male"

CDS

/issue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
17..2641

CDS

/notes="unnamed protein product; hypothetical p-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOP|52540, evidence: SCOP) putative"
/codon_start=1
/protein_id="BAC27438.1"
/db_xref="GI:26327389"

CDS

translation="MAAGRRVGPSPRPRTMAPWKKRLKRRTGASQGRSDSDGGE
FEIQAEDDARARKLGPALPSPTSCVSDVEPTREMVAAQNKKKKSGGFSQMG
SYVPFKGIMKGYKVPPIQKRTIPVLDGKVAVAMTSGKTACFLPMPFERKAR
SACTGARALILSPRELALQMTKTELKFTGLKTAJLIGDQVDEQFAALHFNKPD
IIATPGRLVHVAVEMNLKQSVYVDFEADRLFEMGFEOLOELIGLPGHQTFLF
SATIPKLIVLPAQAGLTPVILRLDVDSKLNQKTSFLVREDTKAVALLYLQNVV
RPQOQTVVFAVKHAEYLTLLMGQVSCAHYISALDQATARKINLAKFTPKSCSTLI
VTDLAAGLDIPLDLVNINYSFPAGKGLFHRVGRVARAGSGTAYSLVAFVYLL
DLHLFLGRSVTLARCEPESVADAGVLDGLRVQSVDDSDSLQTAGASLDLQ
LHRVANAQQOQVRSRPAPESEIKRAKELDLAELGLHPLPSSCFEGEQLRLVDS
IKNVRTTTFEINASSKDPSSOMRAKRODRKAVASFOORROEROGEPADPAORE
LPOREEMVETVSGVTEVVQKRPREGSQGAKRRMETRQDRQRYVYVRKPED
SERGLSVSGACGAEQVAGVLDLMDGDAQNSRGGQQLKWRDKRRKRFVQCSQDEK
KKIKTEGSRFSSIKDLYQKWKQKIDDRSEEGPSNQRGPRGRRGRSQG
TSQPRASSVPAGRMRSSEKTKTEQLKQRRQAKQKORFLQRRGLKQLSARNRRRAQELRQ
GAFGRGAPSRKGRKRM"

FEATURES

polya_signal
2954..2959
/note="putative"

source

2972
/note="putative"

ORIGIN

Alignment Scores:
Pred. No.: 3.47 Length: 2972
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservative: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match: 3.57% Indels: 201
DB: 11 Gaps: 29

ORIGIN

US-09-930-591-2 (1-686) x AK031534 (1-2972)
QY 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114
DB 90 TCGGAACCGCGAAGTCTCCCAAGCCGCGACGAGCTCGATGACGCGAGT 149

ORIGIN

QY 115 IleProValArgArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
DB 150 TCGAGATCCAGCGGAGGATGACGCCGCGGAGAGAGTGGCGCTGGCAGACCTTCG 209

ORIGIN

QY 132 Pro-IleSerTyrLeuLys-----GlySerSerGlyGlyProLeuLe 145
DB 210 CCTCATTTCCATCTCAGCTGCGTATCAGATGTGGAGCCGACACATCGGAGATGTGTC 269

ORIGIN

QY 145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159
DB 270 GAGCCAGAACAGAAAAAGAGAGTCTGGAGGCTTCAGTCCATGGGCTGAGTTACC 329

ORIGIN

QY 160 -CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGlySerLeuGluThr 179
DB 330 CTGTGTTCAAGGGATCATGAAAAAGGGGTACAAAGTGGCGACGCCCATCCAGAGGA 389

ORIGIN

QY 179 hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyr 199
DB 390 CCATC-----CCCGTGATCTTGATGGCAAGATGTGGTGGCCATG----- 430

ORIGIN

QY 199 lnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
DB 431 -----GCCCGACAGGCGAGTGGCAAGAGCGGCTGCTCTCTCTCCCGA 473

ORIGIN

QY 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn 231
DB 474 TGTTTGAGCGGCTGAAGCGACGCGAGTGCAGACGGGGGCTCGAGGCCCTCATCTCTCAC 533

ORIGIN

QY 231 rCsr-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis 248
DB 534 CCACCGGAGAGTGGCCCTCGACACCATGAAGTTCCTAAAGAGCTAGGCAAGTT-CACC 592

ORIGIN

QY 248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrThrGly 264
DB 593 GGCCTCAAGACTGCCTTGCCTCTGGTGGGAGCAAAATGGAAGACCATTTGACGCCCTG 652

ORIGIN

QY 264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGly 284

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9422, URU: <http://genome.gsc.riken.go.jp/>).

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

source

```

142204
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:6030486C03"
/db_xref="MGI:2394945"
/db_xref="taxon:10090"
/clone="6030486C03"
/sex="male"
/tissue_type="testis"
/clone.lib="R1XEN full-length enriched mouse cDNA library"
/dev_stages="13 days embryo"
25. .2754

```

CDS

/note="unnamed protein product; hypothetical p-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOP)52540, evidence: SCOP putative"

```

/codon start=1

```

/protein_id="BAC27509.1"
/db_xref="GI:26327531"
/translation="MAAGRRVGPQPPSPPTMAPWKKKRLKRTGASQGRSDSDDDGE
FEIQGADPARAKLIGPGRALPSEFPFSEUCSDVDPTREMVRAQNKKKKSGGFGMGL
SYPVFGKMKGYKVPTPIQKRTIPEVLDDGVAMARTSGKTAQFLPMPERLAKR
SAQGAARLLISFRETALQMTKFKELGTGLTALITGGQWEDQPAALHENDPI
IITATGRHVHVAFTMLKQSYEVYFDEARLFEMGFPAEQIIGRUPGCHGTVLFI
RATDLPFVARAGLTPVLIRLMDGSKLANQLTFLLDVETAKAALLVLLQNVV
SPDQTVVFATKHAHYLTLLMDGQVSKALNLTALDQTAQKLNKTKCSPLI
VTLAARGLDIPILDVINYSPAGKGLFLHVRGVARGSGTASVILVADPEVPVLL
DHLFLGSLVTLARPCEPSEVADAVGRGVYGRVQSVVDDSDSLQTANGASLDLQGS
LHRVANNAAQQVYFSRRAPSPESIKRAKELLDAELGLPLPFCSEEGEQLKRLVLDS
IKNYRTITTFINASSKDPSSQMRKAKQMDKAVASFPQRRQERQEPADPAPORE
LPQREEMVETVEGTFVYVQKQPRPGSGAKRRMETRQOEYVYRPFQFI
SEGLSVSGAGGAEQVAGVAILDMQDQAKNMSQQQLKWDKRRKFFVGSQBDK
KKIKTSGRTIISSYKEDLYQWQKQIDRDSSEEGPSNQRRGPKRRVGSQBDK
TSQPSGSLVAGNRMLRSKTEQLIKRRQGSAAALPAARGPEAFATSTPSSGPAAP
GRLPQCSLQGGQDEEKDVRSQTPQWGLFVAPGVDRDVHVHLLSDSTGAAL"

polyA signal

```

[{"id": 1, "note": "putative"}]

```

polyA_site 2984

ORIGIN

Alignment Scores:

Pred. NO.:	3.5	Length:	2984
Score:	129.00	Matches:	134
Percent Similarity:	31.84%	Conservative:	55
Best Local Similarity:	22.15%	Mismatches:	216
Query Match:	3.57%	Indels:	201
DB:	11	Gaps:	29

US-09-930-591-2 (1-686) x AK031679 (1-2984)

QY 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114

1096 ATGGGCCAGGGTGTGAGTTGGGCCACATCTATAGTGCCTTGACCAGACG----- 1146

polyA_signal putative"
2960..2965 /note="putative"
polyA_site 2986
/note="putative"
ORIGIN

Alignment Scores:

Pred. No.: 3.5 Length: 2986
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservativity: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match: 3.57% Indels: 201
DB: 11 Gaps: 29

US-09-930-591-2 (1-686) x AK028274 (1-2986)

```

QY 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
94 TCGGAACGCGGAACCTGGGCTTCCCAAGCCGCGACAGCGACTCGGATGACGGCGAGT 153
QY 115 IleProValArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 TCGAGATCCAGGGCGAGGATGACGCCCGCGGAGGAGCTGGGCCCTGGCAGAGCTTGC 213
QY 132 Pro-IleSerTyrLeuLys-----GlySerSerGlyGlyProLeuLe 145
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 CCTATTCTTACTCTCAGAGTGCGTATCAGATGTGGAGCCCGACACTCGGAGATGGTGC 273
QY 145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
274 GAGCCAGACAGAAAGAAAGAGTCTGGAGCTTCCAGTCCATGGGCTGAGTTACC 333
QY 160 -CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGlySerLeuGluThr 179
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
334 CTGTGTTCAAGGGGATCATGAAAGGGCTACAAAGGTGCGCGCCGATCCAGAGGAAGA 393
QY 179 hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrG 199
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
394 CCATC-----CCCGTATCTTGATGGCAAGGATGTGGGCCATG----- 434
QY 199 lnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
435 -----GCCCGACAGGAGTGGCAAGAGCGGCTCTCTCTCCCTCCCGA 477
QY 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnP 231
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 TGTGTAGCGGCTGAAGGACGACGAGTGCACAGACGGGGCTCGAGCCCTCATCTCTCAC 537
QY 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisG 248
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
538 CCACCCGGAGCTGGCCCTGCACACCATGAGTTCTAAGAGCTAGGCALAGTT-CACC 596
QY 248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlyS 264
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
597 GGCTCAAGACTGCTCTGTATCTCTGGTGGGACAAAATGGAACACAGTTGCGAGCCCTG 656
QY 264 exProlIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyA 284
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657 CAGAGAACCTGCACATAATCATGTCACCCCTGGCGCTCTGGTGCATGT-----GGCT 710
QY 284 laTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
711 GTGGAGATGAACCTTGAAGTGCAGAGTGT-----GGAGTATGTGGTG 752
QY 304 le----- 304
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
753 TTGATGAACGACAGAGGCTCTTTGAAATGGGCTTTGCTGACGAGCTACAGGAGATCATA 812
QY 305 -----GlyThrValLeuAsp-----GlnAlaGluThrAlaG 315
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
813 GGCGGCTTCTCTGGGGCCACAGAGCGTGTCTGTCTCTAGCTACACTGCCCGGAGCTGCTG 872

```

```

QY 315 lyAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
873 GTGGAATTTGCACGGCAGCGCTCACAGAGCCCGTCTCATCCGCTGCAGCTAGACTCC 932
QY 332 alProHisProAsnIleGluValAlaLeuSerThrThrGlyGluIlePro----- 349
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
933 AAGCTCAATGACGAGCTCAAGACTC-----CTTCTCTCTGTG 971
QY 350 -----PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyG 363
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
972 CGGAAGACACCAAGGCTGCTGTCTCTCTACCTGCTGCAGAAATGCTGTTGGCCCCAG 1031
QY 363 ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAlaLysLeu 382
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1032 GACCAGCTGTGGTGTCTAGCCACAAGACCATGGAGTACTCTACAGATTGCTG 1091
QY 383 ValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1092 ATGGGCCAGGGTGTGAGTTGCGCCACATCTATAGTGTGCTTGGACACAGC----- 1142
QY 403 ThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1143 -----GCCGCAAGATCAACTTTGGCCCAAGTTACACACAAAC 1178
QY 423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1179 -----AAATGTCCACCTCATCGTGAAGTACCTGCGCGCCGCGGCGCTG 1223
QY 442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1224 GATATCCCACTGCTCGCAACAGCTCATCAACTACAGCTTCCCTCCCAAGGCAAGCTTTC 1283
QY 457 -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIle 473
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1284 CTGCACCGAGTGGCGGTGTGGCCGAGCAGCGCAAGTGGCAGACGC----- 1331
QY 474 TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerSerValLeuCy 493
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1332 TATCTTTGTGGTGGCCCGCAGAGGTCCTTACC----- 1365
QY 493 sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrThrVala 513
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1366 -----TGCTTGACCTACACTGTTCTGGCGCGCTGTGCACCTGGCGCTCTGTGTGAG 1421
QY 513 g-----LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1422 GAGCCTTCAGTGGCAGATGCGTGGCAGG-----ACGAGTG 1460
QY 532 sTyrGluGly----- 535
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1461 CTGGTCCGCTGCCCGAGGTGTAGTGATGATGAGGACAGCAGCTGCAGACTGCCATG 1520
QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1521 GGGGCATCCCTGGATCTTTCAGGGCTGCACC----- 1551
QY 551 rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1552 -----GGTGGCCCAACACGCTCAGCAGCAGTATGTGGCTCA 1589
QY 571 gAlaGlnAlaPro-----ProProSerTyrAsp---GlnMetTr 583
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1590 CGGCCAGCGCTCGCTGAGTCCATCAAGAGACCGCAAGAGCTGGACCTGGCAGAGCTG 1649
QY 583 pLysCys 585
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1650 GGCTTG 1656

```

RESULT 12

BC043699

LOCUS

DEFINITION

BC043699 3956 bp mRNA linear HTC 13-JAN-2003

Mus musculus, Similar to RIKEN cDNA 2410015A15 gene, clone

IMAGE:5344158, mRNA.

```

ACCESSION      BC043699
VERSION        BC043699.1
KEYWORDS       GI:27696772
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE          1 (bases 1 to 3956)
JOURNAL        Strausberg, R.
REMARK        Direct submission
COMMENT        Submitted (10-JAN-2003) National Institutes of Health, Mammalian
                Gene Collection (MGC), Cancer Genomics Office, National Cancer
                Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
                USA
                NIH-MGC Project URL: http://mgc.nci.nih.gov
                Contact: MGC help desk
                Email: cgabbs@mail.nih.gov
                Tissue Procurement: Jeffrey Green M.D.
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Genome Sequence Centre,
                BC Cancer Agency, Vancouver, BC, Canada
                info@bcgsc.bc.ca
                Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
                Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
                Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
                Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
                Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
                Schain, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott,
                Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
                George Yang, Scott Zuyderduyn, Marco Marra.
                Clone distribution: MGC clone distribution information can be found
                through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
                Series: IRAX Plate: 86 Row: 1 Column: 19
                This clone was selected for full length sequencing because it
                passed the following selection criteria: Hexamer frequency ORF
                analysis
                This clone has the following problem: frame shifted.
FEATURES       Location/Qualifiers
                1..3956
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:5344158"
                /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
                ductal carcinoma. 5 month old virgin mouse."
                /clone_lib="NCI CGAP Mams"
                /lab_host="DH10B"
                /note="Vector: pCMV-SPORT6"
ORIGIN
Alignment Scores:
Pred. No.:      5.74      Length:      3956
Score:          129.00    Matches:     134
Percent Similarity: 31.24%  Conservative: 55
Best Local Similarity: 22.15%  Mismatches: 216
Query Match:     3.57%    Indels:      201
DB:              11       Gaps:         29
US-09-930-591-2 (1-686) x BC043699 (1-3956)
QY      100 CysGlySerAspLeuTyLeu-----ValThrArgHisAlaAspVal 114
Db      80 TCGGAAGACCGCGAGTCTCCCAAGCCGCGACGACTCGGATGACGGCGAGT 139
QY      115 IleProValArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
Db      140 TCGAGATCCAGCGGAGGATGACGCCCGGGGAGGAGTGGCCCTGGCAGAGCCTTGC 199
QY      132 Pro-IleSerTyLeuLeu-----GlySerSerGlyGlyProLeuLe 145
Db      115 -----

```

```

QY 442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db 1210 GACATCCCACTGCTGGCAACAAGTCATCAACTACAGCTTCCTGCCAAGGCAAGCTCTTC 1269
QY 457 -----ValSerArgThrGlnArgArgGlyValArgThrGlyArgGlyLysProGlyIle 473
Db 1270 CTCACCGAGTGGCGTGTGGCCCGCAGCAGCGCGGAGGTGGCACACC----- 1317
QY 474 TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerValLeuCy 493
Db 1318 TATCTCTTGGTGGCCCGCAGCAGAGTCCCTACC----- 1351
QY 493 sGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThrThValAr 513
Db 1352 -----TGCCTGACCTACACCTGTTCCTGGCCCGCTCTGTACCCCTGGCCCGCTCTGTGAG 1407
QY 513 g-----LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
Db 1408 GAGCCTTCAGTGCAGATCGGTGGCAGGG-----ACGAGGTG 1446
QY 532 eTrpGluGly----- 535
Db 1447 CTGGTGGCTGCCCGCAGAGTGTAGTGGATGATGAGGACAGCAGCCTGCAGACTGCCATG 1506
QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
Db 1507 GGGGATCCCTGATCTTACGGCCCTGCACC----- 1537
QY 551 rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
Db 1538 -----CGTGGCCCAACACAGCTCAGCAGCAGTATGTGGCTCA 1575
QY 571 galaginAlaPro-----ProProSerTrpAsp---GlnMetTr 583
Db 1576 CGGCCAGCGCCCTCGCTGATGCCATCAAGAGAGCCAGAGCTGACCTGCAGAGCTG 1635
QY 583 pLysCys 585
Db 1636 GGCTTGC 1642

RESULT 13
AY404177
LOCUS
DEFINITION
1697 bp DNA linear GSS 15-DEC-2003
Mus musculus DDX34 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY404177
VERSION
AY404177.1 GI:39760154
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1697)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Infering nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1697)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers

```

```

source
1. 1697
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1.->1697
/gene="DDX34"
/locus_tag="HCM1803"

gene

ORIGIN

Alignment Scores:
Pred. No.: 1.44 Length: 1697
Score: 128.50 Matches: 134
Percent Similarity: 35.56% Conservative: 68
Best Local Similarity: 23.59% Mismatches: 221
Query Match: 3.53% Indels: 148
DB: 29 Gaps: 30

US-09-930-591-2 (1-686) x AY404177 (1-1697)
QY 52 ValCysTrpThrValTyrHisGlyValAlaGlyThrArgThrIleAlaSerProLysGlyPro 71
Db 116 GTCCTCTTGAAGAATGATATTCTTCCGCG---ATGAGGATTATATCC-----GTC 160
QY 72 ValIleGln-----MetTyrThrAsnValAspGlnAspLeuVal 84
Db 161 GTGGTTCAGAGGAGTGTGCAGAAATTCTGGGCTTCTTTGAACGCTGCAGAGATTCGAGC 220
QY 85 GlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCysGlySerSerAsp 104
Db 221 ACCTCAAGACCTCCACAGAAGAAGAGGACCCCTGGCATGCCCAACATGGCATCGCTG 280
QY 105 Leu-----TyrLeu---ValThrArgHisAlaAspValIleProValArgArg 120
Db 281 CTTTAGCTGACCTACCTCTCACTTACGACCCACGCT----- 316
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 317 ---ACCGCATCAACCTCTCCATTCTCAGCCAGACACTCGGGGTCGTCATGGGCCCGGCA 373
QY 141 GlyGly-----ProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAla 158
Db 374 GGGGGTCCCCCGAGAGAGTGTCCGAGTCC---GCCGTGCTTCTCTCCTACCTACCTAG 430
QY 159 ValCysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeu---Glu 177
Db 431 ACTTCCAGCAGA-----AGCAGGCTTTTGGAGGCTCGCCAAACTGCAGCGTGAGCGAG 484
QY 178 Thr-ThrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGln-S 197
Db 485 CGGCACCTTCCCATCGCCAGATGGGAACCGCAT-----CCTGCAGACTCTCAAGG 535
QY 197 eTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAla 217
Db 536 AACACCAAGTGGTGGTGGCGCGGACACAGAGCTGTGGCAAGTCTACTCAGGTGCCCC 595
QY 217 laAlaTyrAlaAlaGlnGlyTyrLys---ValLeuValLeuAsnPro----- 231
Db 596 AGTACTTGTCTGGCCCGCGCTTCTCAGTCAGCGGCATGCATCAGCCCGGAGAAATCGCT 655
QY 232 -----SerValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleA 250
Db 656 GCATCTCACTGCGCAAGCGGTGGCTTCGAGAGCCTCAGTCAGTATGGCTCCCGAGTTG 715
QY 250 spProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerT 270
Db 716 GCTACCAAGATCCGCTTT---GAAAGCCTCGGTGGCAGGACCAACAGATCGTGTCTCTGA 772
QY 270 hrTyrGlyLysPheLeu-----AlaAspGlyGlyCysSerGlyValAlaTyrAspIleI 288
Db 773 CGGTGGGCTCTCTCTACGCCAGATCAGCGGTGAGCCAGCTCCCGCAGTACCAAGTCC 832
QY 288 leIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValL 308

```

```

Db      833 TGATCGTGGACGAGGTCCACGAGCGACCTGCACATGACTTCTCTGCTGGGT----- 885
Qy      308 euAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThrAlaThrProG 328
Db      886 -----GTCTGCGACGCGCTGCTGCC----- 906
Qy      328 lySerValThrValProHisProAsnIleGluValAlaLeuSerThrThrGlyGluI 348
Db      907 -----CAGCGGCTGACCTCAAGTCTATCTCTGCTGGCCACCATCAACA 952
Qy      348 leProPheTyrGlyLysAlaIleProLeuGluAlaIleLysGlyGlyArgHisLeuIleP 368
Db      953 TCTCTCTCTCTCCAGC-----TACT 973
Qy      368 heCysHisSerLysLysLysCysAspGluLeuAlaIleLysLeuValAlaLeuGlyValA 388
Db      974 TCAGCCATGCT-----CCGGTGGTGCAGGTGCTGGAGGCTCTTCCCATCAGCGTGT 1027
Qy      388 snAlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValV 408
Db      1028 TCGATGTGGCT-----CCAGCTGGAGTCCGGAAGT 1057
Qy      408 alValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp----PheAspSerValI 427
Db      1058 GCATCTCTCCACCAACATTGCCGAGACCTCAGTCAGATCGATCGATCGCTTCTGCTGG 1117
Qy      427 leAspCysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrI 447
Db      1118 TGGATCTGGG-----AAGGTGAAGAGATGAGTACGACCCACAGGCCCAAC 1165
Qy      447 leGluThrIleThrLeuProGlnAsp-----AlaValSerArgThrGlnArgA 463
Db      1166 TGCAGCGCTG-----CAGAAATCTGGATCAGCAGGCGCATGCTGACGACGCA 1216
Qy      463 rgGlyArgThrGlyArgGlyLysProGlyIle---TyrArgPheValAlaProGlyGluA 482
Db      1217 AGGCCGAGCGGCGCGACGCGTCCCGGCTGCTACCGTCTCTATGCGGATCTGAC- 1275
Qy      482 rgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCysAlat 502
Db      1276 -----TATGATGCTT-TGGCC- 1292
Qy      502 rpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThrProG 522
Db      1293 -----GTACCAGTCCCG 1305
Qy      522 lyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeuThrH 542
Db      1306 GAATCCG-----CAGATGAA-GAGTATGAGTGTGGGGACCCACG 1346
Qy      542 isIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuV 562
Db      1347 ACCTTCCCTTCATCGAGCCCTCCACCTGCGAGCGTGGAMACCGTATCTCT-ATCT 1405
Qy      562 alaIaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProPro-----ProSerTrpA 580
Db      1406 CCAAGACAGCGGGCCCTGGACAGCTGGAGGCGCTCACCCCATCAGGTTCCTGCTGCG 1465
Qy      580 spGlnMetTrpLysCysLeu 586
Db      1466 CC-----AGCTGCTG 1476

RESULT 14
BU054966
LOCUS
DEFINITION UT-M-FD0-b2i-n-19-0-UI.r1 NIH BMAP_FDO Mus musculus cDNA clone
ACCESSION BU054966
VERSION BU054966.1
KEYWORDS BU054966.1 GI:22495043
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 691)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..691
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6404178"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5 dpc"
 /lab_hosts="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FDO"
 /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is TCAGAGAGCC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 0.369 Length: 691
 Score: 127.50 Matches: 45
 Percent Similarity: 44.19% Conservative: 31
 Best Local Similarity: 26.16% Mismatches: 69
 Query Match: 3.52% Indels: 27
 DB: 13 Gaps: 7
 US-09-930-591-2 (1-686) x BU054966 (1-691)
 Qy 174 GluSerLeuGluThrThrMetArgSer---ProValPheSerAspAsnSerSerProPro 192
 Db 145 GAATCTATCCAGGTGTGCCGCCGAGCTGCTCCCTGTTCCCTTCCGAGGAGGACTTCTG 204
 Qy 193 AlaValProGlnSerTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSer 212
 Db 205 GCCGCCATTGCCACCATCAGTCTCTCATCATCGAAGCGAGACTGGCTCTGGGAAGACC 264
 Qy 213 ThrLysValPro-----AlaAlaTyrAlaAlaGlnGlyTyrLysValLeu 227
 Db 265 ACACAGATCCACAGTACCTCTTTGAGAGCGGTACACAAAGAGAGGCGCATCAGATTGCT 324
 Qy 228 ValLeuAsnProSerValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis 247
 Db 325 TGCACCCAGCCCCGAGAGTGGCGGCTATGAGTGTGCACGCCGAGTGGCCGGAGATG 384
 Qy 248 Gly-----IleAspProAsnIleArg-----ThrGlyVal 257

Db 385 GGTGTGAAGCTTGGGAACGAGGTGGCTACAGATCCGGTTTGAGACTGCACCTCAGAG 444

QY 258 ArgThrIleThrThrGlySerProIleThrTyrSerThrTyrGly-----Lys 273
 |||||:::|
 Db 445 CGAAGTGT-----CTCCGTACATCAGACAGATGGAATGCTACCTCCGAGAG 489

QY 274 PheLeuAlaaspGlyGlyCysSerGlyGlyAlaTyrAspIleIleCysaspGluCys 293
 |||||:::|
 Db 490 TTCCTCTCTGAGCTGACCTTGCA-----AGTTACAGTGTGTGATGTGATGAAGCT 543

QY 294 HisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThr 313
 |||||:::|
 Db 544 CACAGCGGACCTTGACACAGACATCTCTTGGATTGATCAAGACGTCGCTAGATTC 603

QY 314 AlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
 |||||:::|
 Db 604 CGACCTGAGCTCAAGGTCCTGTGGTTCAGGCACA 639

RESULT 15

CD240900 738 bp mRNA linear EST 22-MAY-2003

LOCUS AGENCOURT_14124904 NIH_MGC_177 Mus musculus cDNA clone

DEFINITION IMAGE:30380772 5', mRNA sequence.

ACCESSION CD240900

VERSION CD240900.1 GI:31001364

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 738)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Michael Brownstein Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDCM173 row: f column: 13
 High quality sequence stop: 482.

FEATURES
 Location/Qualifiers
 1..738
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30380772"
 /lab_host="DH10B (Ti-phage-resistant)"
 /clone_lib="NIH_MGC_177"
 /note="Organ: liver; Vector: pDNR-LIB; Site_1: SfiI (ggcattatggcc); Site_2: SfiI (ggccgctcgcc); cDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
 5'-AAGCAGTGGTATCAGCGAGGTGGCATTAGCGCGGG-3' and
 5'-ATTCTAGGCGGAGCGCGACATG-gt(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5 kb size fraction. Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	0.414	Length:	738
Score:	127.50	Matches:	45
Percent Similarity:	44.19%	Conservative:	31
Best Local Similarity:	26.16%	Mismatches:	69
Query Match:	3.52%	Indels:	27
DB:	14	Gaps:	7

US-09-930-591-2 (1-686) x CD240900 (1-738)

QY 174 GluSerLeuGluThrThrMetArgSer---ProValPheSerAspAsnSerSerProPro 192
 |||||:::|
 Db 62 GAATCTATCCAGGCTGTGCGCGCAGCGCTGCCGCTGTTCCCTTCCGAGAGGAGCTTCTG 121

QY 193 AlaValProGlnSerTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSer 212
 |||||:::|
 Db 122 GCCGCCATTGCAACACCATCAGGTCTCTCATCTCGAAGCGCAGACTGGCTCTGGGAAGACC 181

QY 213 ThrLysValPro-----AlaAlaTyrAlaAlaGlnGlyTyrLysValLeu 227
 |||||:::|
 Db 192 ACACAGATCCACAGTACCTCTTTGAGGAGGGTTACACAAGAAGGCGCATGAAGATTGCT 241

QY 228 ValLeuAsnProSerValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis 247
 |||||:::|
 Db 242 TGCACCCCGCGCGAGAGTGGCGCTATGAGTGTGCAGCCCGAGTGGCGCGGAGATG 301

QY 248 Gly-----IleAspProAsnIleArg-----ThrGlyVal 257
 |||||:::|
 Db 302 GGTGTGAAGCTTGGGAACGAGGTGGCTACAGCATCCGGTTTGAGGACTGCACCTCAGAG 361

QY 258 ArgThrIleThrThrGlySerProIleThrTyrSerThrTyrGly-----Lys 273
 |||||:::|
 Db 362 CGAACTGTT-----CTCGCTACATGACAGATGGAATGCTACTCCGAGAG 406

QY 274 PheLeuAlaaspGlyGlyCysSerGlyGlyAlaTyrAspIleIleCysaspGluCys 293
 |||||:::|
 Db 407 TTCCTCTCTGAGCTGACCTTGCA-----AGTTACAGTGTGGTGGTGGATGAAGCT 460

QY 294 HisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThr 313
 |||||:::|
 Db 461 CACGAGCGGACCTTGCACACAGACATCTCTTTGGATTGATCAAGACGTCGCTAGATTC 520

QY 314 AlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
 |||||:::|
 Db 521 CGACCTGAGCTCAAAAGTCTCTGGTGGCTTCAGGCACA 556

Search completed: September 16, 2004, 00:40:41

Job time : 4609 secs

This Page Blank (uspio)